BMJ Open

Correct diagnosis of childhood pneumonia in public facilities in Tanzania: a randomised comparison of diagnostic methods

Taylor Salisbury,[1] Alice Redfern,[2] Erin K Fletcher ,[1] Jean Arkedis,[1] Felix Bundala,[3] Alison Connor,[2] Ntuli A Kapologwe,[4] Julius Massaga,[5] Naibu Mkongwa,[3] Balowa Musa,[6] Cammie Lee[1]

To cite: Salisbury T, Redfern A, Fletcher EK, *et al*. Correct diagnosis of childhood pneumonia in public facilities in Tanzania: a randomised comparison of diagnostic methods. *BMJ Open* 2021;**11**:e042895. doi:10.1136/bmjopen-2020-042895

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-042895).

Received 23 July 2020
Accepted 07 May 2021

For numbered affiliations see end of article.

Correspondence to
Dr Erin K Fletcher;
efletcher@r4d.org

ABSTRACT

Objective This study compares two methods for clinical diagnosis of childhood pneumonia that aim to estimate rates of underdiagnosis and overdiagnosis of childhood pneumonia by examining the sensitivity of Integrated Management of Childhood Diseases implementation in routine care against lung ultrasound (LUS) diagnosis.

Setting We conducted observations in 83 public health facilities (dispensaries, health centres and district hospitals) in Pwani, Dodoma and Tabora, Tanzania between October and December 2017.

Methods We used a novel method to estimate rates of underdiagnosis and overdiagnosis of childhood pneumonia by comparing directly observed public provider diagnoses to the results of diagnoses made by trained clinicians using Mindray DP-10 ultrasound machines. We perform multivariate analysis to identify confounding effects and robustness checks to bound the result. We also explore a number of observable characteristics correlated with higher rates of agreement between provider diagnoses and ultrasound diagnoses.

Results We observed 93 providers conducting exams on patients aged 2 months–5 years who presented respiratory symptoms or were given a respiratory diagnosis by the provider. Of these 957 patients, 110 were excluded from analysis resulting in a final sample of 847.

17.6% of cases identified as pneumonia via LUS examinations in our sample were diagnosed as pneumonia by providers, suggesting that a significant number of pneumonia cases for which care is sought in the public sector go undiagnosed. Provider knowledge of breath counting and years of experience are positively correlated with higher agreement. While clinical examination rates are not statistically correlated with agreement, it is notable that providers conducted a clinical examination on only about one-third of patients in the sample.

Conclusion Our results suggest that provider training and knowledge of clinical examination protocols for pneumonia diagnosis are predictive of correct diagnosis of pneumonia and should be further explored in future research as a tool for improving quality of care.

Strengths and limitations of this study

- This is the first study in Tanzania to estimate rates of overdiagnosis and underdiagnosis of childhood pneumonia using lung ultrasound and to present broadly generalisable estimates of these metrics.
- This study goes beyond measurement of overdiagnosis and underdiagnosis of childhood pneumonia and explores the facility-level and provider-level characteristics that are associated with diagnosis outcomes.
- Severe cases of pneumonia that were immediately identified and referred for higher-level care are not included, and so these findings are most applicable and generalisable to non-severe cases of pneumonia.
- This study was conducted on a select sample of children whose caregivers sought treatment in the public system, which introduces significant bias in estimating or projecting the prevalence of pneumonia and mortality risk on a population level.

INTRODUCTION

Pneumonia is the single largest infectious cause of death among children under 5 globally.[1] In Tanzania, pneumonia morbidity and mortality have decreased substantially in the past decade and a half, but rates of childhood pneumonia remain among the highest in the world. Seemingly at odds with stalled progress, there is broad consensus on efficacious treatments for childhood pneumonia[1] and global standards for diagnosis.[2 3] Oral amoxicillin, preferably in dispersible tablet (DT) form (amox DT), has been designated the first-line treatment for children under 5 by the WHO[4] and adopted by the government of Tanzania as the standard of care.[5] Low availability of this first-line treatment remains a problem in Tanzania; only about two-thirds of facilities had the treatment available in

the most recent national Service Provision Assessment in 2014.

Alongside access to first-line treatments, accurate and appropriate diagnosis and prescription, and care-seeking are key to reducing pneumonia-related mortality. Evidence around the quality of provider care and accuracy in diagnosing pneumonia is thin. A recent pilot in Malawi suggests that providers vastly underdiagnose childhood pneumonia,[6] but there is little evidence on the level of misdiagnosis, or the directionality of it, for childhood pneumonia in Tanzania or in other settings. If correct diagnosis rates are low, the potential gains from improving diagnosis are large as care-seeking rates for acute respiratory infections are relatively high at 70%.[7]

The recommended diagnostic for identifying pneumonia in children is clinical observation, though this method is highly dependent on provider ability. In ambiguous or contested cases, chest radiography (CXR) is recommended,[8] though even in high-resource settings CXR should be combined with other diagnostic methods to ensure correct diagnosis.[9] CXR is an expensive tool requiring sophisticated equipment and the presence of skilled radiologists,[8] and in low-resource settings, such as Tanzania, barriers to implementing CXR are high. The Tanzanian Ministry of Health, Community Development, Gender, Elderly, and Children (MoHCDGEC) has adopted the WHO's Integrated Management of Childhood Diseases' (IMCI) guidelines as the national standard of care.[6] IMCI outlines three steps to determine a pneumonia diagnosis in children under the age of 5[6] and has been shown to have a sensitivity of 69.9% and a specificity of 59.6% when compared against point-of-care ultrasound imaging.[10] Despite work by the government of Tanzania and partners to encourage full adoption of, and adherence to, IMCI, it is unclear how often clinicians attempt to use the protocol in practice and whether they adhere to it sufficiently well to make appropriate diagnoses. If IMCI is poorly understood, retained or administered, there could be high rates of either overdiagnosis or underdiagnosis of childhood pneumonia.[11–13]

In recent years, ultrasound radiography has been applied to diagnose a widening array of diseases.[14] Specifically, there is mounting evidence that lung ultrasound (LUS) can be used to accurately diagnose pneumonia in children under 5, with multiple studies showing that LUS is nearly as, or more, accurate than CXR.[15–17] A meta-analysis across eight studies including 765 children found LUS to have a sensitivity of 96% and a specificity of 93%[9] when compared with CXR. Additionally, international evidence-based recommendations suggest that LUS is particularly suitable for identifying lung consolidations, the principal pathological sign of pneumonia.[18] LUS can be conducted using portable, battery-powered ultrasound machines. As such, LUS provides a more objective approach to estimate the rate of underdiagnosis in rural and remote health facilities in comparison with IMCI as opposed to traditional methods of comparing clinical assessments between observers.

We seek to fill the gap in knowledge of rates of overdiagnosis and underdiagnosis of pneumonia—and the correlates of quality of care and diagnosis—in the public sector in Tanzania by asking: what is the number of diagnosed and undiagnosed cases of childhood pneumonia for which care is sought in the public sector? To do so, we employ a combination of healthcare provider surveys, direct observations and lung ultrasonography of potential cases of pneumonia in Tanzania public health facilities. We seek to estimate the sensitivity of provider diagnosis of childhood pneumonia by identifying potential cases of pneumonia using LUS re-examinations and then comparing those cases to the original provider diagnoses, which are supposed to be based on IMCI. The role of both LUS and IMCI in this study was solely to serve as measurement tools, and we did not seek to demonstrate the efficacy or effectiveness of LUS or IMCI as point-of-care diagnostics.

METHODS

Field methods

Study design and setting

We conducted a prospective, observational study in children under 5 between 23 October and 20 December 2017, across the Pwani (East zone), Dodoma (Central Zone) and Tabora (West Zone) regions in Tanzania. These regions were purposefully selected to create a balanced mix of geographies, rural and urban areas, and high-performing and low-performing regions. The sample of health facilities was selected with multistage random sampling using STATA V.14. We randomly selected three districts per region and then randomly selected health facilities from each of the three tiers of the Tanzanian public healthcare system, weighted by the approximate prevalence of each facility type: one hospital, one health centre and seven dispensaries within each district. In districts without a hospital, an eighth dispensary was randomly selected instead, resulting in a sample of 83 facilities. Enumerators conducted limited stock audits at each selected facility to track the availability of amox DT and collected basic background information about the facility.

Provider selection and observation

The director-in-charge identified all eligible healthcare providers at his or her facility. Providers were eligible for inclusion if they worked full time, were responsible for the assessment and treatment of children under 5, and were available during the observation period. One provider was randomly selected from the list of eligible providers in each facility and was asked for consent to having a study enumerator observe their patient consultations. If the provider did not consent or became unavailable over the course of the observation period, another provider was randomly selected to replace him or her. Prior to observation, enumerators administered a knowledge test to the selected provider.

The enumerators for this study did not have any prior medical qualifications; enumerators were provided the IMCI guidelines for pneumonia and, post-training, were capable of identifying the key steps of the algorithm. During patient consultations, enumerators completed a checklist to record provider actions and adherence to the IMCI guidelines for the assessment of a child with respiratory symptoms. The ICMI guidelines for a patient with respiratory symptoms are outlined in online supplemental table A in the annex. Enumerators observed the routine consultations of the selected healthcare provider over multiple days and referred eligible study participants for diagnosis verification via LUS. Spot checks were conducted to ensure that enumerators were accurately capturing provider activity and to identify any potential provider behaviour change associated with being observed. Providers were asked to record their final diagnosis and treatment plan for every patient regardless of age and diagnosis to mitigate potential observation bias.

Selection of children for secondary assessment

Children aged 2 months–5 years who presented with any respiratory symptoms (cough, difficulty breathing or fast breathing) or were given a respiratory diagnosis by the provider (regardless of symptoms) were eligible to participate in the study. During patient observations, eligible children were identified via at least one of two methods: (1) provider's report of any respiratory symptoms that was provided confidentially to the enumerator or (2) enumerator observation of any respiratory symptoms. Providers gave any non-respiratory diagnoses, prescriptions and instructions to the caregivers or parents and then caregivers or parents of identified children with respiratory symptoms were approached outside the consultation room by the enumerator–observer to request informed consent to secondary assessment via LUS in a separate, private observation area by a trained clinician without the presence of the provider. In the case where the provider did not ask the caregiver about any respiratory symptoms, enumerators posed these questions after the consultation outside of the provider's room. Children were excluded if the provider referred them to a higher-tier facility for immediate treatment for severe illness, including but not limited to severe pneumonia, or if the study ultrasound clinician felt referral was warranted on initial reassessment. This selection process is summarised in figure 1 (957 children in the sample had respiratory symptoms, and 110 of the patients were excluded from analysis for various reasons). Figure 2 in the Results section provides a breakdown of participant inclusion and exclusion.

Care was taken to ensure that the additional diagnostics were conducted in a private space separate from the provider's room. Following the observation period, enumerators conducted a postobservation tool with the provider, and the technicians provided private, constructive feedback to the providers on their adherence to IMCI protocols.

Lung ultrasound

Ultrasound clinician training

Clinicians who had been previously trained on IMCI were identified by MoHCDGEC and given (1) 2 weeks of theoretical and practical training on LUS and (2)

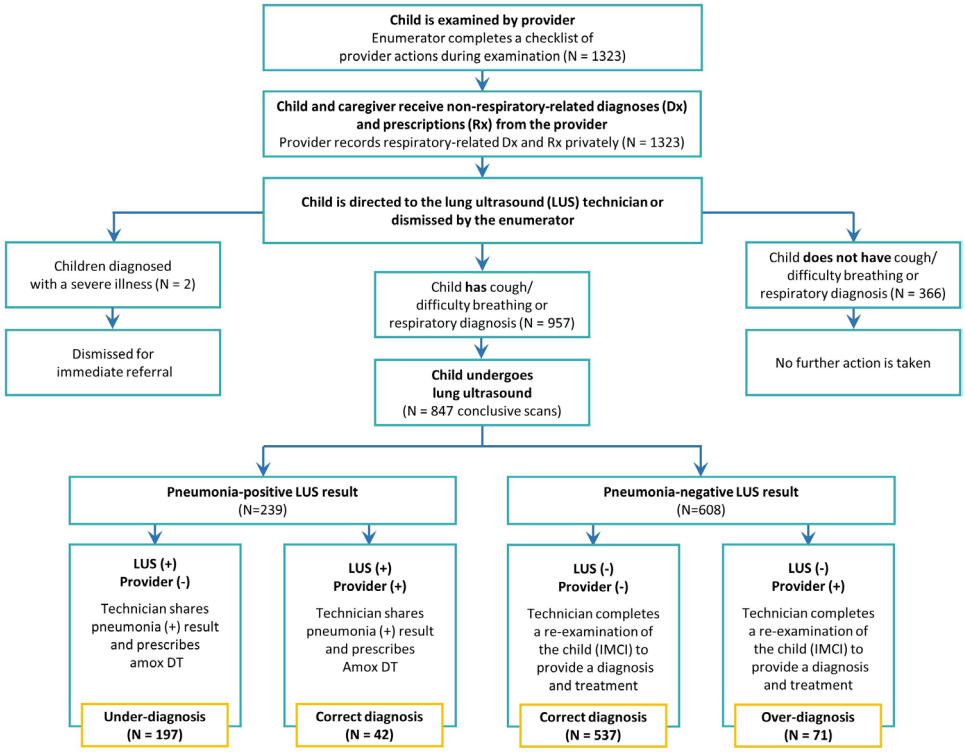

Figure 1 Protocol for identification of correctly diagnosed cases. DT, dispersible tablet; IMCI, Integrated Management of Childhood Diseases

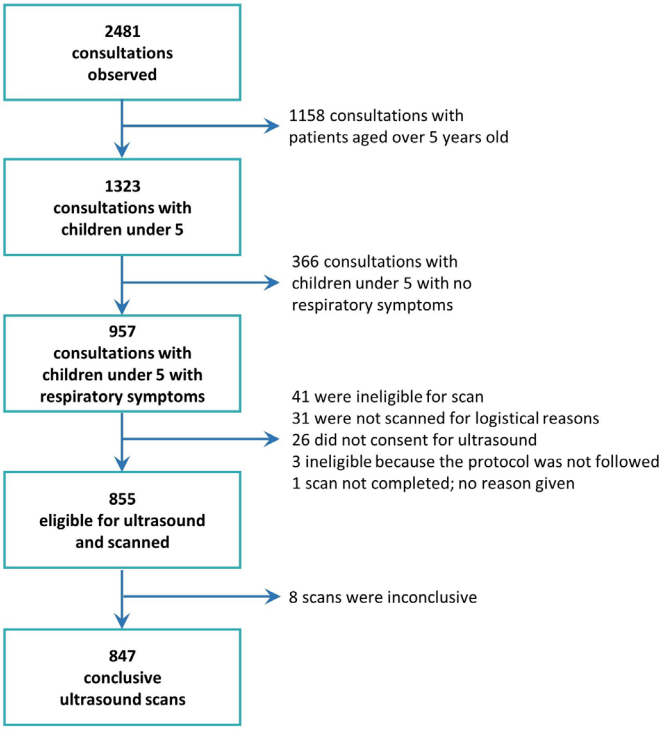

Figure 2 Summary of study participant inclusion and exclusion.

refresher IMCI training by lead radiologists from Muhimbili University of Health and Allied Sciences (MUHAS) and Aga Khan Hospital of Dar es Salaam, followed by close supervision during 2 weeks of piloting. LUS clinicians were contracted to participate in the study if they demonstrated the ability to conduct high-quality scans and a 90% agreement rate with LUS reviewers—a panel of three lead radiologists from MUHAS—based on scans conducted during piloting.

Ultrasound assessment

LUS clinicians captured ultrasound images across eight distinct lung regions in each child and established a diagnosis based on an algorithm developed by the lead radiologists and based on current LUS literature (table 1). On making a pneumonia-positive diagnosis, the LUS clinician informed the caregiver and prescribed and supplied a course of amox DT. If the diagnosis was pneumonia negative, the LUS clinician conducted an IMCI reassessment and consulted the provider's diagnosis to determine the appropriate treatment.

Scan review

A random selection of scans from each LUS clinician was sent for independent review by one of the LUS reviewers. Where there was disagreement between the clinician and reviewer, a second LUS reviewer reviewed the scan. If there was disagreement between the two reviewers, a third LUS reviewer was consulted. Scans where no consensus could be reached by the team of three LUS reviewers were excluded from the final analysis. Results of the random review were monitored to ensure no LUS clinicians fell below a 90% agreement rate with reviewers; extra training, support and additional monitoring was provided if issues were identified.

Data analysis

Main analysis

Data analysis was performed in STATA V.14. We assessed overdiagnosis and underdiagnosis of childhood pneumonia by estimating the percentage of cases identified as pneumonia-positive or pneumonia-negative via LUS and comparing them to the number of cases identified as pneumonia-positive or no diagnosis by providers. The analysis was guided by a preanalysis plan, included in the annex, with two important deviations. First, we conducted the analysis at the case level rather than the provider level, and thus conduct analysis with and without clustering at the provider level. Second, we modified our definitions of overdiagnosis and underdiagnosis to be consistent with the case-level analysis. Here, we examine underdiagnosis only among the pneumonia-positive (via LUS) patients and only look at overdiagnosis among the pneumonia-negative (via LUS) patients.

Multivariate analysis

In order to test for confounding variables and to identify possible reasons for these low rates, we performed

Table 1 Diagnostic criteria used for the diagnosis of pneumonia on lung ultrasound

	Consolidation	B-lines	Effusion
Criteria	At least one region with evidence of consolidation:	At least one region with at least three B-lines in one image (within two rib spaces) but not B-lines across all eight regions, and not B-lines in the last intercostal space of inferior zones only (that may capture the stomach, spleen or liver)	Evidence of pleural effusion:
	Hypoechoic or anechoic area with blurred margins of >1 cm with either		Homogenous, anechoic fluid in the pleural space that is unilateral and mild or moderate (not severe)
	Positive shredding-like sign, or positive air or fluid bronchogram		

multivariate analysis using logistic regression on the sample of LUS scans. We controlled for patient, provider and facility characteristics, with and without clustering at the provider level.

Robustness checks

Acknowledging that our estimate may be subject to survey-induced bias, we employed a number of robustness checks in order to bound the estimate. First, we conducted a 'leave-one-out' analysis by iteratively dropping all of the observations from each of the eight LUS clinicians who conducted the LUS exams. If one LUS clinician was particularly inaccurate at diagnosing pneumonia via ultrasound, their observations may exert an outside influence on the mean.

We also employed a randomised leave-one-out robustness check to assess unobservable bias. We drop a randomly selected 10% of observations and calculated the underdiagnosis rate from the remaining observations, repeating 1000 times. Additionally, we conducted checks to see if LUS clinician diagnosis performed better over time, that is, with practice, and in different geographical regions. Results of the leave-one-out analysis are reported in online supplemental table B in the annex.

Patient and public involvement

This research was done without the involvement of patients and their caregivers. Patients and their caregivers did not contribute to the development of the research question, design and conduct of the study, determination of outcome measurements, recruitment to the study, nor plans for dissemination of the study results.

RESULTS

Between October and December 2017, we observed 93 providers (12% medical attendants, 40% nurses, 5% assistant clinical officers, 23% clinical officers and 19% physician clinicians) with a target sample of 800 observations consisting of 10 per facility over a 5-day period. Providers conducted 2481 consultations with patients of all ages during the study time frame of 23 October–30 December 2017, of which 1323 patients were children under the age of 5. Of these children under 5, 957 presented with any respiratory symptoms, and conclusive ultrasounds were conducted for 847 cases, which represents our analysis sample. A summary of study participant inclusion and exclusion is presented in figure 2. These cases were observed in 83 health facilities in three regions (27 in Dodoma, 20 in Pwani and 26 in Tabora) consisting of 67 dispensaries, 11 health centres and 5 district hospitals in the sample. The distribution of pneumonia-positive cases (as diagnosed by providers and by LUS) is described in table 2. The distribution of pneumonia-positive cases, with a focus on specificity, is described in online supplemental table C in the annex.

Lung ultrasonography was conducted with conclusive diagnosis for 847 children presenting with respiratory systems. All images were reviewed by the study's LUS clinicians on site, and then a subsample was reviewed by a blinded panel of expert independent reviewers. The panel consisted of three lead radiologists from Muhimbili University, and in total they reviewed a total of 260 scans (30% of all scans). There was disagreement on the final diagnosis in 23 scans (9% of those reviewed); in total, 14 of the disagreements (5% of scans reviewed) resulted in a change in the final diagnosis. Ultimately, eight scans were given an inconclusive diagnosis, primarily due to technical problems capturing scan images, leaving 847 scans with conclusive diagnoses which were used to verify provider diagnosis.

Underdiagnosis

Only 17.6% of children presenting respiratory symptoms were diagnosed with pneumonia by both the provider and the LUS clinician (table 3). These results vary by child age; providers were slightly more likely to diagnose pneumonia-positive cases identified as such by LUS in children under 1 year (20.8%) than children aged 1–5 (15.2%).

Overdiagnosis

Only 37.4% of cases diagnosed by providers as pneumonia were confirmed as clinical pneumonia by the LUS clinicians. These results also vary by age, with higher sensitivity among cases of pneumonia in children under 1 year old (table 3). These findings point to significant overdiagnosis of pneumonia as well as the underdiagnosis described previously.

Multivariate analysis

Multivariate analysis yielded several notable results around the effect of provider training. When we control for confounding factors, provider training in IMCI is strongly associated with higher sensitivity; providers who have had any IMCI training are almost nine times more likely to have their pneumonia diagnoses confirmed by LUS compared with providers who have had no training at all. This effect is controlling for any effects that arise from having counted breaths or performing a clinical exam (providers counted breaths for 13% of patients and conducted a clinical exam for 34% of patients who eventually received LUS (see table 4 for more details). Provider role (or cadre) is not predictive of sensitivity of diagnosis, but we note that the sample sizes for each role are small and thus it is difficult to distinguish statistically significant effects. Somewhat counterintuitively, provider years of experience is associated with lower sensitivity. Providers with more experience are less likely to have their diagnosis of pneumonia confirmed by the LUS. Finally, provider knowledge is also important. Knowing the IMCI step of counting breaths for 1 min is associated with three times higher likelihood that a provider will positively diagnose a LUS-confirmed case of pneumonia.

Table 4 shows the abbreviated output of this analysis. Specifications with provider-level clustering are reported

Table 2 Summary of children with a completed ultrasound examination

	(1) Cases with a completed LUS (n)	(2) Total cases with a completed LUS (%)	(3) Provider Dx+ cases (n)	(4) LUS Dx+ cases (n)	(5) Provider Dx+ and LUS Dx+ cases (n)	(6) Pneumo+ according to provider (%)	(7) Pneumo+ according to LUS (%)	(8) Sensitivity†
						As % of full sample (col. 3÷col. 1)	As % of full sample (col. 4÷col. 1)	As % of LUS pneumo+ cases only (col. 5÷col. 4)
Total	847	100.0	113	239	42	13.3	28.2	17.6
Gender								
Male	428	50.5	59	129	23	13.8	30.1	17.8‡
Female	419	49.5	54	110	19	12.9	26.3	17.3
Child age (months)								
2–11	288	34.0	48	101	21	16.7	35.1	20.8‡
12–59	559	66.0	65	138	21	11.6	24.7	15.2
Provider carried out any clinical examination§								
Yes	284	33.5	58	91	27	20.4	32.0	29.7***
No	563	66.5	55	148	15	9.8	26.3	10.1‡
Provider counted child's breaths								
Yes	107	12.6	29	33	15	27.1	30.8	45.5***
No	740	87.4	84	206	27	11.4	27.8	13.1‡
Duration of provider examination (min)								
≤2	155	18.3	18	45	8	11.6	29.0	17.8‡
3–5	403	47.6	45	101	13	11.2	25.1	12.9
6–10	218	25.7	32	71	15	14.7	32.6	21.1
>10	70	8.30	18	22	6	25.7	31.4	27.3
Provider role at facility								
Medical attendant	102	12.0	6	32	4	5.9	31.4	12.5
Nurse	340	40.1	39	101	13	11.5	29.7	12.9
Assistant clinical officer	46	5.4	7	10	3	15.2	21.7	30
Clinical officer	195	23.0	34	50	12	17.4	25.6	24
Physician clinician	164	19.4	27	46	10	16.5	28.0	21.7‡
Years of experience provider has diagnosing and treating children under 5								
≤2	187	22.1	23	44	4	12.3	23.5	9.1‡
3–5	385	45.5	60	111	28	15.6	28.8	25.2**
6–10	120	14.2	14	45	7	11.7	37.5	15.6
11–15	37	4.4	4	7	0	10.8	18.9	0.0

Continued

Table 2 Continued

	(1) Cases with a completed LUS (n)	(2) Total cases with a completed LUS (%)	(3) Provider Dx+ cases (n)	(4) LUS Dx+ cases (n)	(5) Provider Dx+ and LUS Dx+ cases (n)	(6) Pneumo+ according to provider (%)	(7) Pneumo+ according to LUS (%)	(8) Sensitivity†
						As % of full sample (col. 3÷col. 1)	As % of full sample (col. 4÷col. 1)	As % of LUS pneumo+ cases only (col. 5÷col. 4)
16–20	52	6.1	8	12	1	15.4	23.1	8.3
>20	66	7.8	4	20	2	6.1	30.3	10.0
Received IMCI training								
In the past 2 years	136	16.1	20	46	14	14.7	33.8	30.4***
2 or more years ago	74	8.7	12	13	4	16.2	17.6	30.8*
No training received	622	73.4	78	175	22	12.5	28.1	12.6‡
Provider knows to count child's breaths								
Yes	563	66.9	95	146	34	16.9	25.9	23.3***
No	279	33.1	18	92	8	6.5	33.0	8.7‡
Month when examination took place								
October	139	16.4	21	58	13	15.1	41.7	22.4‡
November	455	53.7	69	113	21	15.2	24.8	18.6
December	253	29.9	23	68	8	9.1	26.9	11.8
Region								
Pwani	282	33.3	51	102	24	18.1	36.2	23.5‡
Dodoma	297	35.1	35	73	8	11.8	24.6	11.0**
Tabora	268	31.6	27	64	10	10.1	23.9	15.6
Facility type								
Dispensary	678	80.0	88	205	34	13.0	30.2	16.6‡
Health centre	118	13.9	18	29	7	15.3	24.6	24.1
Hospital	51	6.0	7	5	1	13.7	9.8	20.0

Significance compared with the reference group denoted as follows: *p<0.1, **p<0.05, ***p<0.01.
†Reference group for pairwise significance tests indicated by ‡.
§Clinical examination refers to the completion of one or more of the three assessment steps providers should be undertaking when diagnosing a child under five according to the IMCI algorithm: (1) counting breaths, (2) looking for chest in-drawing, and/or (3) listen to breathing for sounds of stridor or wheezing.
Dx, diagnosis; IMCI, Integrated Management of Childhood Diseases; LUS, lung ultrasound.

Table 3 Sensitivity of provider pneumonia diagnosis, by age of child

	LUS Dx+	LUS Dx–	Total
A. All children 2–59 months			
Provider Dx+	42	71	113
Provider Dx–	197	537	734
Total	239	608	847
	Sensitivity (17.6%)	Specificity (88.3%)	
B. Children 12–59 months			
Provider Dx+	21	44	65
Provider Dx–	117	377	494
Total	138	421	559
	Sensitivity (15.2%)	Specificity (89.5%)	
C. Children 2–11 months			
Provider Dx+	21	27	48
Provider Dx–	80	160	240
Total	101	187	288
	Sensitivity (20.8%)	Specificity (85.6%)	

Dx, diagnosis.

in online supplemental table D in the annex; specifications with additional controls for the chest region where pneumonia was identified and enumerator and LUS clinician fixed effects are available on request. Due to the high number of dichotomous variables, we also performed these regressions using a linear probability model. Though the coefficients are by nature not comparable in magnitude to ORs estimated by a logistic model, levels of significance are consistent across the specifications.

Robustness checks results

We employed a series of robustness checks in order to bound the initial sensitivity estimate of 17.6%, acknowledging potential biases that could arise from various sources. Excluding observations for any individual LUS clinician changes the estimate of underdiagnosis by a maximum of 2.8 percentage points above or below the mean. Using this method, we bound the estimate at 79.6%–85.2% of cases going undiagnosed. We cannot disentangle these effects from any influence the study enumerators may have had on LUS outcomes as each LUS clinician was paired with only one enumerator throughout the data collection process.

We also used a simulation method to bound the estimate by dropping 10% of the observations and recalculating sensitivity. We repeated this 1000 times. With the output from the simulation, we bound our estimate using the 5th–95thpercentiles of simulations, resulting in an estimated underdiagnosis rate of 81.1%–83.9%. These estimates are robust to dropping a randomly selected 1%–20% of observations per simulation.

Finally, we bound our estimate using the sensitivity of lung ultrasonography, estimated at 96%.[8] Tables with robustness checks are available in the annex. We are confident that our system of double review reduced incidence of misdiagnosis that was a result of LUS clinician error, but if we account for the sensitivity of the instrument, we estimate that the rate of underdiagnosis would change by less than 1 percentage point in either direction.

We found no evidence that LUS clinician diagnosis was more prone to error in early weeks of analysis than later weeks or in particular areas over and above the bounds noted earlier for pairs of enumerators and clinicians.

DISCUSSION

Main findings

To our knowledge, this is the first study that uses LUS as a standard of reference for estimating rates of overdiagnosis and underdiagnosis of childhood pneumonia in Tanzania. Our findings point to both significant overdiagnosis and underdiagnosis of childhood pneumonia. We estimate that a mere 17.6% of cases where care is sought in the public system in our mixed purposive and randomised, non-nationally representative sample are correctly diagnosed as pneumonia (as verified by the LUS reference standard). We propose bounding of the estimate based on robustness checks to conclude that between 14.8% and 19.4% of childhood pneumonia cases are correctly diagnosed in Tanzania. Alongside, overdiagnosis is also common; we estimate that 62.8% of cases of pneumonia diagnosed by providers in our sample are in fact not clinical pneumonia.

Our multivariate analysis suggests that providers who were trained recently were more likely to diagnose pneumonia in line with the LUS clinicians. While we cannot make causal claims around the effectiveness of training, other work, such as the Lancet Global Health Commission, has shown the modest positive effects of training on diagnosis rates.[19] However, much of the literature on training effectiveness is based on studies with small sample sizes conducted over short time horizons. In view of the findings from this study, we suggest that future work focus on developing and evaluating effective strategies to improve diagnosis and the quality of clinical care.

Limitations

Our data allow us to conduct a directed, specific evaluation of LUS as a reference standard for diagnosing childhood pneumonia in Tanzania and to estimate overdiagnosis and underdiagnosis rates of a high-burden infectious disease. Our sample has several key limitations. First, the sample includes three regions of Tanzania we chose as being reflective of national-level trends but does not constitute a nationally representative sample. Thus, we cannot estimate with great precision the rate of overdiagnosis and underdiagnosis for the whole country and

Table 4 Multivariate analysis using logistic regression, clustering at provider level

	Provider diagnosis			LUS diagnosis			Sensitivity of provider diagnosis (n=234)			Specificity of provider diagnosis (n=597)			Provider and LUS positive (n=831)		
	Pneumonia-positive (n=831)			Pneumonia-positive (n=831)											
	OR (95% CI)		P value	OR (95% CI)		P value	OR (95% CI)		P value	OR (95% CI)		P value	OR (95% CI)		P value
Clinical exam conducted															
Yes	1.25	(0.58 to 2.68)	0.567	1.40	(0.85 to 2.30)	0.185	3.00	(0.69 to 12.95)	0.142	0.91	(0.42 to 1.96)	0.806	2.28	(0.74 to 7.01)	0.149
No	1			1			1			1			1		
Provider counted child's breaths															
Yes	2.32	(1.03 to 5.26)**	0.043	0.86	(0.43 to 1.73)	0.678	2.79	(0.83 to 9.30)*	0.096	0.62	(0.24 to 1.55)	0.305	2.32	(0.78 to 6.91)	0.131
No	1			1			1			1			1		
Exam duration (min) (log)	1.10	(0.60 to 2.00)	0.766	1.08	(0.78 to 1.49)	0.643	0.85	(0.46 to 1.56)	0.596	0.64	(0.31 to 1.35)	0.245	0.87	(0.50 to 1.51)	0.621
Provider experience (years) (log)	0.62	(0.35 to 1.09)*	0.099	1.41	(0.98 to 2.02)*	0.064	**0.43**	**(0.24 to 0.75)*** **	**0.003**	1.44	(0.76 to 2.72)	0.260	0.72	(0.43 to 1.23)	0.230
Provider IMCI training															
Received training in the past 2 years	1.39	(0.46 to 4.21)	0.564	1.76	(0.79 to 3.91)	0.165	**8.92**	**(2.22 to 35.84)*** **	**0.002**	2.66	(0.59 to 12.01)	0.203	**7.39**	**(2.23 to 24.46)*** **	**0.001**
Received training 2+ years ago	1.59	(0.48 to 5.30)	0.447	0.58	(0.21 to 1.55)	0.276	**14.63**	**(2.50 to 85.71)*** **	**0.003**	1.24	(0.31 to 5.05)	0.761	5.00	(0.96 to 26.19)*	0.057
None	1			1			1			1			1		
Provider knows to count child's breaths															
Yes	2.72	(0.94 to 7.81)*	0.064	0.51	(0.27 to 0.96)**	0.038	3.07	(0.83 to 11.40)*	0.093	0.29	(0.08 to 1.03)*	0.056	1.31	(0.50 to 3.47)	0.586
No	1			1			1			1			1		
Controls															
Child age (decimal years)	0.88	(0.74 to 1.05)	0.155	0.82	(0.70 to 0.96)**	0.011	0.90	(0.57 to 1.43)	0.654	1.05	(0.84 to 1.31)	0.662	0.73	(0.50 to 1.06)	0.100
Child is female															
Yes	0.95	(0.64 to 1.41)	0.782	0.80	(0.57 to 1.13)	0.208	0.91	(0.36 to 2.26)	0.831	1.06	(0.68 to 1.65)	0.792	0.78	(0.35 to 1.74)	0.549
No	1			1			1			1			1		

The dependent variable in each estimation is indicated in the column header; significance is denoted as follows: *p<0.1, **p<0.05, ***p<0.01.
IMCI, Integrated Management of Childhood Diseases; LUS, lung ultrasound.

the rate of underdiagnosis may actually be much higher than estimated here if there is significant regional variation in the practice or quality of care by providers. Additionally, we acknowledge that there is significant potential bias in estimating or projecting the prevalence of pneumonia and mortality risk on a population level from the select sample of children whose parents seek out care for them. Caregivers who seek treatment in the public system may be different from those who seek care in the private system and may be different from those who decline to seek care at all. Key characteristics that we do not observe include, but are not limited to, geographical dispersion, risk aversion, socioeconomic status or children's observed symptoms or disease progression that could influence the rate at which pneumonia is correctly diagnosed. Finally, we excluded cases where children presented with severe illnesses that required referral to a higher-tier facility. These cases were few, but included, though were not limited to, severe pneumonia. As such, we may be underestimating or overestimating the rate of correct diagnosis of all pneumonias, given that there was no comparison diagnostic applied to an unknown number of severe pneumonia cases.

That we did not seek to specifically measure the value of training but see it as a potentially important component to improving quality of care underscores the need for randomised control trials and further testing of IMCI trainings and support. However, it is clear from our results that appropriate diagnosis is unlikely to occur when physical exams are not routinely conducted: low rates of agreement between LUS and provider diagnosis are underscored by the low adherence to IMCI protocols—only 13% of cases were observed with providers counting breaths and 34% included any physical exam.

However, there are clear gains to be made in improving diagnosis and access to frontline treatment, as evidenced by the low rates of correct diagnosis. As noted previously, many other studies have sought to prove the efficacy and effectiveness of LUS as a point-of-care diagnostic. This was beyond the scope of this study.

An additional possible limitation of the study is that the presence of observers may have induced potential Hawthorne effects, altering healthcare workers' clinical diagnosis. In the case that there were Hawthorne effects, we would expect them to bias our results positively, suggesting even more extreme underdiagnosis than we estimate. As such, we could consider these results an underestimate. However, the direction and magnitude of Hawthorne effects are difficult to ascertain, and thus we implemented several safeguards and checks to mitigate against Hawthorne effects, and we are convinced by our efforts. In a comparison of rates of pneumonia diagnosis during our study period to a period 2 weeks early by comparing registers showed no significant differences, suggesting that there was a little to no impact on provider behaviour.

In implementation, we endeavoured to mitigate Hawthorne effects by using established methods of observation in sub-Saharan Africa as opposed to methods such as cameras or provider questionnaires. We then built in appropriate checks to the study design. First, observations were conducted by enumerators who were not medically trained, so that any explicit or implicit signalling was avoided. These enumerators were researchers, only trained to understand how to correctly identify assessment steps in the IMCI protocol. Second, enumerators observed all of the clinical providers' consultations throughout their multiple days at the health facility to mitigate the providers changing behaviours for a specific age group or set of symptoms. Third, we conducted the pneumonia diagnosis and treatment questionnaire to assess provider knowledge only at the end of the data collection period to further mitigate any potential signalling. We originally planned to exclude the first 10 observations of each provider, based on a study conducted in Tanzania that found direct observation improving providers' practice for an initial number of cases before the Hawthorne effects wore off.[20] However, during piloting, we found that providers were rarely completing the IMCI assessment steps even with the enumerator present and thus did not exclude the observations.

Other evidence and implications

Mounting evidence shows that misdiagnosis of pneumonia has significant potential to do harm. Overdiagnosis or incorrect diagnosis increases the risk of global antibiotic resistance.[3] Underdiagnosis increases the risk of mortality and morbidity.[7] Our estimate of underdiagnosis of childhood pneumonia in Tanzania is close to estimates recently derived for Malawi,[7] suggesting the need for a global assessment to establish whether this is a wider phenomenon. Moreover, the differential correct diagnosis rates we observe based on age indicate that there may be a need for modifications to the IMCI protocol or trainings on treating children as they age.[11]

CONCLUSION

Our findings point to significant underdiagnosis of pneumonia in Tanzania and the subsequent conclusion that many children at risk are not receiving life-saving drugs. Our findings are consistent with research from Malawi, where correct diagnosis was also low[7] and where evidence shows that IMCI adherence is low.[11] Overall, it is clear that despite recent decreases in childhood pneumonia mortality, we may quickly reach the limit of improvements in the global burden of disease if provider diagnosis does not improve.[3]

Author affiliations
[1]Results for Development Institute, Washington, DC, USA
[2]IDinsight, Nairobi, Kenya
[3]Reproductive and Child Health Section, Ministry of Health, Community Development, Gender, Elderly, and Children, Dodoma, Tanzania
[4]Social Welfare and Nutrition Services, President's Office Regional Administration and Local Government, Dodoma, Tanzania
[5]National Institute for Medical Research, Dar es Salaam, Tanzania

[6]Department of Radiology and Imaging, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Dar es Salaam, Tanzania

Acknowledgements The study was funded by Good Ventures. EKF, TS, JA, and CL are employees of Results for Development Institute, which has received funding for work on this from Good Ventures. AC and AR are employees of IDinsight. FB and NM are employees of the Tanzania Ministry of Health, Community Development, Gender, Elderly, and Children. NAK is an employee of the President's Office, Regional Administration and Local Government. JM was an employee of the National Institute of Medical Research at the time of this research. BM is an employee of Muhimbili University of Health and Allied Sciences.

Contributors EKF drafted and edited the manuscript. AR, JA and CL drafted sections of the manuscript. TS, JA, CL, EKF and AC edited the manuscript. TS, AR and EKF performed the statistical analysis. FB, NAK, JM, NM and BM contributed to study design and assisted in the implementation of the protocols. CL, JA, AC and AR contributed to the study design. TS and AR participated in and monitored the study implementation.

Funding This work was funded by Good Ventures. The funder does not issue award/grant numbers.

Competing interests Results for Development reports grants from Good Ventures and the Bill and Melinda Gates Foundation on pneumonia in East Africa.

Patient consent for publication Not required.

Ethics approval This study was approved by the Tanzania Commission for Science and Technology and the National Institute for Medical Research (NIMR, reference number NIMR/HQ/R.8a/Vol. IX/2561). The participants gave informed consent before taking part in the study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The full dataset and statistical code used in this analysis are available from the corresponding author (efletcher@r4d.org). Data are anonymised, with all personally identifiable information removed, therefore posing low risk of identification.

ORCID iD

Erin K Fletcher http://orcid.org/0000-0002-2804-6890

REFERENCES

1 World Health Organization. Pneumonia fact sheet, 2016. Available: http://www.who.int/news-room/fact-sheets/detail/pneumonia [Accessed Dec 2019].
2 Walker CLF, Rudan I, Liu L, *et al*. Global burden of childhood pneumonia and diarrhoea. *Lancet* 2013;381:1405–16.
3 International Vaccine Access Center (IVAC), Johns Hopkins Bloomberg School of Public Health. Pneumonia and diarrhea progress report 2018, 2018. Available: https://www.jhsph.edu/ivac/resources/2018-pneumonia-and-diarrhea-progress-report/ [Accessed April 2019].
4 Unicef Supply Division. Amoxicillin dispersible tablets (dT): product profile, availability and guidance, 2013. Available: https://www.unicef.org/supply/files/Amoxicillin_DT_Product_Profile_and_Supply_Update.pdf [Accessed April 2019].
5 World Health Organization. *Handbook IMCI: integrated management of childhood illness*. Geneva, Switzerland: World Health Organization, 2000.
6 Uwemedimo OT, Lewis TP, Essien EA, *et al*. Distribution and determinants of pneumonia diagnosis using integrated management of childhood illness guidelines: a nationally representative study in Malawi. *BMJ Glob Health* 2018;3:e000506.
7 Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS) and ICF. *2015-16 Tanzania demographic and health survey and malaria indicator survey (TDHS-MIS) key findings*. Dar es Salaam, Tanzania and Rockville, Maryland, USA: MoHCDGEC, MoH, NBS, OCGS, and ICF, 2016. https://dhsprogram.com/pubs/pdf/SR233/SR233.pdf
8 Pereda MA, Chavez MA, Hooper-Miele CC, *et al*. Lung ultrasound for the diagnosis of pneumonia in children: a meta-analysis. *Pediatrics* 2015;135:714–22.
9 Htun TP, Sun Y, Chua HL, *et al*. Clinical features for diagnosis of pneumonia among adults in primary care setting: a systematic and meta-review. *Sci Rep* 2019;9:7600.
10 Chavez MA, Naithani N, Gilman RH, *et al*. Agreement between the world Health organization algorithm and lung consolidation identified using point-of-care ultrasound for the diagnosis of childhood pneumonia by general practitioners. *Lung* 2015;193:531–8.
11 Bjornstad E, Preidis GA, Lufesi N, *et al*. Determining the quality of IMCI pneumonia care in Malawian children. *Paediatr Int Child Health* 2014;34:29–36.
12 Lange S, Mwisongo A, Mæstad O. Why don't clinicians adhere more consistently to guidelines for the Integrated Management of Childhood Illness (IMCI)? *Soc Sci Med* 2014;104:56–63.
13 Armstrong Schellenberg J, Bryce J, de Savigny D, *et al*. The effect of integrated management of childhood illness on observed quality of care of under-fives in rural Tanzania. *Health Policy Plan* 2004;19:1–10.
14 Schaub B, Gueneret M, Jolivet E, *et al*. Ultrasound imaging for identification of cerebral damage in congenital Zika virus syndrome: a case series. *Lancet Child Adolesc Health* 2017;1:45–55.
15 Samson F, Gorostiza I, González A, *et al*. Prospective evaluation of clinical lung ultrasonography in the diagnosis of community-acquired pneumonia in a pediatric emergency department. *Eur J Emerg Med* 2018;25:65–70.
16 Jones BP, Tay ET, Elikashvili I, *et al*. Feasibility and safety of substituting lung ultrasonography for chest radiography when diagnosing pneumonia in children: a randomized controlled trial. *Chest* 2016;150:131–8.
17 Boursiani C, Tsolia M, Koumanidou C, *et al*. Lung ultrasound as first-line examination for the diagnosis of community-acquired pneumonia in children. *Pediatr Emerg Care* 2017;33:62–6.
18 Volpicelli G, Elbarbary M, Blaivas M, *et al*. International evidence-based recommendations for point-of-care lung ultrasound. *Intensive Care Med* 2012;38:577–91.
19 Kruk ME, Gage AD, Arsenault C, *et al*. High-quality health systems in the sustainable development goals era: time for a revolution. *Lancet Glob Health* 2018;6:e1196–252.
20 Leonard K, Masatu MC. Outpatient process quality evaluation and the Hawthorne effect. *Soc Sci Med* 2006;63:2330–40.