Vaccine 38 (2020) 6618–6626

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

An assessment of the quality of vaccination data produced through smart paper technology in The Gambia

Alieu Sowe *, Maria Isabella Gariboldi

MyChild Solution External Project Evaluators, Gambia

ARTICLE INFO

Article history:
Received 24 March 2020
Received in revised form 28 July 2020
Accepted 31 July 2020
Available online 8 August 2020

Keywords:
Smart paper technology
Electronic immunisation register
Data quality
Vaccination
eHealth

ABSTRACT

Introduction: MyChild Solution is an innovative Electronic Immunisation Register (EIR) reliant on Smart Paper Technology, thereby eliminating the need for electronic devices and internet connectivity at the point-of-care. The goal of this study is to characterise the quality of routine immunisation data generated using MyChild Solution compared to data obtained through the conventional health management information system (HMIS) used in The Gambia.

Method: We used the World Health Organization's (WHO) Data Quality Review (DQR) Toolkit to evaluate MyChild Solution's data quality in the 19 health facilities across two regions implementing MyChild Solution in The Gambia at the time of the evaluation. We evaluated all applicable data quality metrics as well as additional metrics of interest, including the incidence of recording errors, the incidence of incomplete indicator level data, and implausible dates. Where possible, we compared results to those of the conventional HMIS.

Results: Both MyChild Solution and the conventional HMIS produced 100% complete and timely data in their reference years. Both systems had no moderate or extreme outliers and showed the expected Penta 1 to Penta 3 dropout direction. However, the proportion of verification factors that are not acceptable was higher in the conventional HMIS. MyChild Solution was found to near perfectly (99.98%) digitise scanned documents. These and other data quality indicators evaluated demonstrate that MyChild Solution produces high quality data with high completeness, timeliness, and consistency compared to the conventional HMIS system.

Conclusion: MyChild Solution produces high quality data as per the DQR Toolkit metrics and other metrics of interest of interest. The more internally consitent data produced through MyChild Solution compared to the conventional HMIS demonstrates its potential for supporting data-driven decision-making in immunisation.

1. Introduction

Health information systems are recognized by the World Health Organization (WHO) as one of the key building blocks of health systems, their main functions being data generation, compilation, analysis and synthesis, and communication and use [1]. These functions support all other health system building blocks and form the bedrock for sound decision making within and across the building blocks. Electronic health information systems could be crucial to improving population health in low- and middle-income countries given their potential to provide real-time individual-level data, thereby enabling comprehensive monitoring and a better understanding of health outcomes of interest compared to traditional paper-based systems [2]. In traditional paper-based systems, data is not available in real-time especially at higher levels of the health system and individual-level records are only available at service delivery points, as the data is aggregated as it moves up the data flow hierarchy. This compromises the quality of data available to decision-makers, as well as poses issues at the point-of-care, as individual records are not as easily searchable or retrievable in the case of patients moving between health facilities.

One of the health service domains quickly attracting the use of electronic health information systems is immunisation programs. Electronic Immunisation Registers (EIRs) improve access, availability, and quality of vaccination data [3]. Improving vaccination data quality in low- and middle-income countries is a necessary step to sustaining and improving vaccination gains [4]. Good quality vaccination data promotes accountability [4] and provides a more

* Corresponding author.
E-mail address: sowea1@yahoo.com (A. Sowe).

https://doi.org/10.1016/j.vaccine.2020.07.074
0264-410X/

accurate picture of areas that are not meeting immunisation targets, which in turn is useful for planning interventions to improve vaccination coverage. The use of EIRs has the potential of improving vaccination programs as it provides enhanced monitoring capabilities, including tracking individual vaccination schedules. However, ensuring their successful introduction, adoption, and sustainability requires very careful prior planning, taking into account several issues including funding, sustainability, and capacity building [5–7]. This necessitates substantial financial investments compared to traditional paper-based systems. The Pan American Health Organization (PAHO) has developed guidelines on practical considerations to consider from the planning to the evaluation of EIRs [8]. Apart from financial, equipment (e.g. computers and tablets), and human resources required for implementation, some EIRs have additional requirements that constrain their successful implementation in low-resource settings. For example, EIRs based on electronic data collection systems at the point-of-care are vulnerable to internet connectivity and electricity issues, rendering them inappropriate solutions in communities with unreliable internet access and electricity. An EIR called MyChild Solution was introduced in The Gambia by Shifo Foundation and its partners in 2017. MyChild Solution is an innovative health management information system (HMIS) for immunisation based on Smart Paper Technology (SPT), optimised for use in low-resource settings and enabling the maintenance of an EIR. SPT relies on customised paper forms that are designed to be automatically digitised after being scanned using optical character recognition (OCR) technology, machine learning algorithms, and built-in data validation and recognition rules. This allows sustaining a digital registry while maintaining the reliability and low cost of paper at the point-of-care, thus ensuring that patient data collection will not be affected by power outages or connectivity issues during immunisation sessions. In The Gambia, at the time of the evaluation, MyChild Solution was in place in 19 clinics across Western Region 1 (WR1) and Western Region 2 (WR2) for use in the Expanded Programme on Immunisation (EPI) services. SPT forms scanning and data verification processes occur at the regional health directorates (RHDs) for these regions.

Fig. 1 summarises the basic processes involved in MyChild Solution of relevance to this assessment. A more comprehensive overview of MyChild Solution and SPT can be found at Shifo Foundation's website (https://shifo.org/en/solution/). Infants are registered in the system through Birth Registration forms. The registration box for each child in these forms contains a unique identification number called a MyChild ID. The immunisation, vitamin A and deworming services they receive are tallied onto the Health Records forms along with the child's MyChild ID. At the end of each month, health workers also fill out a Vaccine Management and Data for Action form, in which they indicate relevant information for logistics management, including the number of vaccines received, vaccine stock balances, and cold chain equipment information. Examples of the forms used by implementing health facilities can be found in the Supplementary Material.

In the conventional system in place before MyChild Solution, health workers would register infants in paper-based immunisation registers, and would subsequently manually update doses received in immunisation registers and tally books. Completing monthly return forms required a lengthy process involving manually extracting vaccine doses consumed using tally books, stock balances from the previous month, and vaccine doses received during the current month.

The introduction of MyChild Solution radically simplified this process as the system automatically generates health facility Monthly Return forms, requiring only the additional input of vaccine stock numbers (received and balance) at the end of the month. After scanning the forms, a manual data quality assurance process is used to resolve any fields on forms that are flagged by the system. Fields might be flagged due to issues such as poor handwriting or due to non-compliance with validation rules. Where possible, flagged fields are resolved manually by regional staff by comparing unrecognised fields to original scanned forms. Additional validation rules are applied to the immunisation registry, for example, to estimate birth dates based on immunisation doses received for children with missing or unreasonable birth dates. MyChild Solution does not estimate dates of birth for children who are registered without one but haven't been tallied as receiving any immunisation, deworming or vitamin A service as their

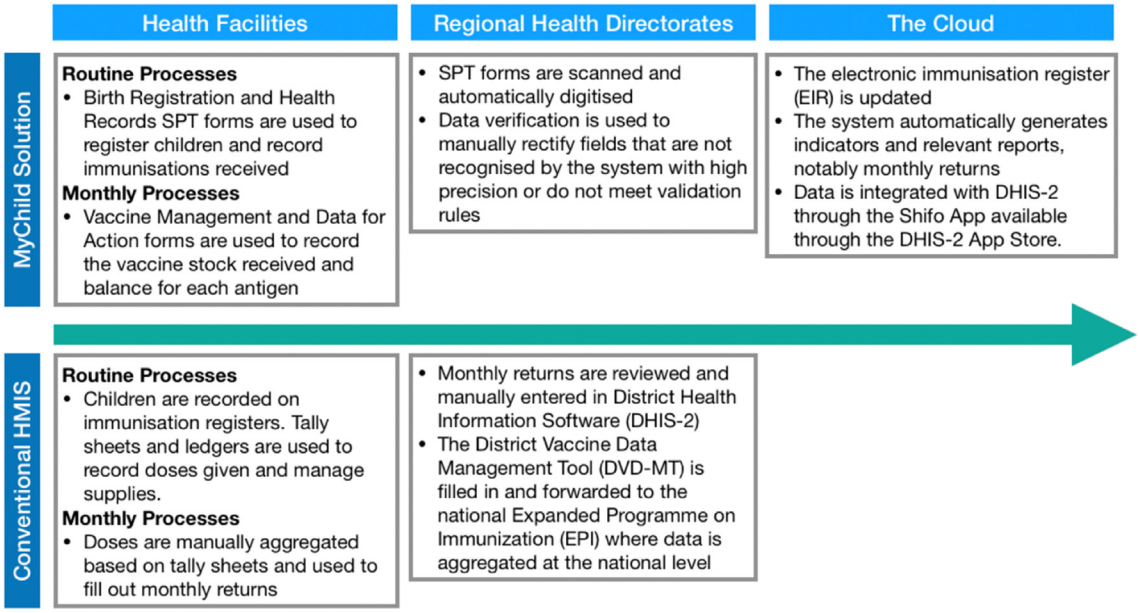

Fig. 1. Summary of relevant MyChild Solution and the conventional HMIS processes. Please note that only processes relevant to understanding the data quality review undertaken are included for simplicity.

date of birth cannot be estimated. The system also randomly assigns children with a missing sex. The estimation of dates of birth and the random assignment of children to a sex in the event of missing information is to facilitate the aggregation of monthly summaries, which are done by sex and age categories. A child's information in the EIR can be updated in case of missing or erroneous information using the MyChild Birth Records Update form.

MyChild Solution produces several outputs. A set of 16 Key Performance Indicators (KPIs) including data quality, coverage of selected antigens, missed opportunities for vaccination, dropout and wastage rates, and vaccination timeliness are sent to health workers through SMS at the end of each month for them to update their paper KPI monitoring graphs. This empowers health workers to identify actions to improve their own performance based on their KPI trends. The system also generates a defaulters list for each health facility at the end of every month to facilitate defaulter tracing. Vaccine needs for the next month are also automatically estimated by the solution and used to inform logistics management.

Scope of the evaluation

The most basic requirement for an HMIS to be useful in routine activities and decision-making, is to guarantee the provision of high-quality data. This evaluation provides a thorough assessment of data quality of both MyChild Solution and the conventional HMIS based on the Data Quality Review (DQR) Toolkit (2017) [9]. This toolkit was developed by the WHO, The Global Fund, Gavi, and the United States Agency for International Development (USAID)/MEASURE Evaluation in 2017. Its purpose is to guide periodic independent assessments of data that is reported at the facility level and is available through health information systems. For example, Bhattacharya et al. (2019) applied this framework to assess the quality of data for maternal and newborn indicator monitoring in the District Health Information Software 2 (DHIS-2) [10]. Various adaptations were made to the DQR methodology, when required, based on the specific scale and nature of the project, as detailed in the methodology section. Further metrics of interest were evaluated to provide a more comprehensive characterisation of MyChild Solution. The evaluation aimed to address the following specific questions:

1. Is there any difference in data quality between data produced through MyChild Solution and data produced by the conventional system based on the WHO DQR Toolkit?
2. What is the level of consistency between vaccines recorded in MyChild Health Records forms and the EIR?
3. What is the level of data completeness of immunisation sessions captured in MyChild Solution?
4. Is there any difference in data quality indicators between data collected at fixed and outreach sessions in MyChild Solution?
5. How common are likely recording errors, incomplete indicator level data, and implausible dates in MyChild Solution?

2. Materials and methods

2.1. WHO DQR Toolkit

The WHO DQR Toolkit promotes a unified approach to assessing reported data quality from the health facility to the national level [9]. It proposes several data quality indicators (metrics) that are grouped into four dimensions, namely: (1) the completeness and timeliness of data; (2) the internal consistency of data; (3) the external consistency of data, i.e. its agreement with other sources of data such as surveys, and (4) the external comparisons of population data, i.e. a review of denominator data used to calculate rates for performance indicators [9]. These metrics were applied to the data as outlined in Table 1. Additional details can be found in the WHO DQR Toolkit.

Unless specified in Table 1, all data quality indicators under the DQR Toolkit were evaluated using data over a 12-month period (March 2018 - February 2019) and included all the health facilities using MyChild Solution in the two regions. For the conventional HMIS, the most recent 12 months before the introduction of MyChild Solution were used.

2.2. Other data quality indicators

The data quality evaluation also included indicators of interest beyond the DQR Toolkit's framework. These are:

1. **Incidence of recording errors:** We checked two types of recording errors using electronic register data spanning March 2018 to February 2019:

1. **Proportion of children recorded as having received multiple doses of Bacillus Calmette–Guérin (BCG) vaccine during different visits:** The motivation for choosing BCG is that a child is expected to be vaccinated with BCG vaccine only once and the chance of administering the same vaccine twice will be lower than that for other antigens, such as Hepatitis, because BCG forms a scar when properly administered.
2. **Proportion of children recorded as having received multiple doses of the same vaccine during the same visit:** This error was checked for Oral Polio vaccine (OPV), Pentavalent vaccine, Rotavirus vaccine, Pneumococcal conjugate vaccine, and Measles-containing vaccine because a child is required to receive multiple doses of each one of these vaccines before completing the immunisation schedule. percentage.

2. **Comparison between fixed and outreach immunisation sessions:** Data quality between fixed and outreach immunisation sessions was compared using two indicators:

1. **Completeness of immunisation sessions captured in MyChild Solution for fixed and outreach immunisation sessions:** This indicator was measured by comparing the number of fixed and outreach vaccination sessions captured in MyChild Solution's dashboard against the monthly planned fixed and outreach vaccination sessions by health facilities.
2. **Incidence of recording errors in fixed and outreach immunisation sessions:** This was evaluated by disaggregating each of the two types of recording errors assessed (multiple doses of BCG during different visits and multiple doses of the same vaccine during the same visit) by fixed and outreach immunisation session types.

3. **Incomplete data entry and implausible dates:** We defined incomplete data entry as a recorded variable for which some observations have no data (empty), such as missing dates of birth or sex. We defined implausible dates as dates of birth that are unreasonable as they would presume a child is too old to receive immunization services (older than five years of age at registration) or is born in the future relative to the recorded session date (which could reflect both an erroneous immunization session date or an erroneous date of birth). The reason for choosing an age limit of 5 years is that this is the age limit for the national EPI program for routine childhood vaccinations.

2.3. Data analysis

Data analysis was carried out using Microsoft Excel, Stata [11] and R [12].

Table 1
Summary of the DQR Toolkit application for the evaluation of data quality of MyChild Solution in The Gambia.

Metric	Definition	How the metric was evaluated
Dimension 1: Completeness and Timeliness of Data		
1.1 Completeness of facility reporting	Percentage of expected facility monthly reports (previous 1 year) that are actually received.	March 2018 to February 2019 monthly reports of all health facilities implementing MyChild Solution were used. * The last 12 months before the implementation of MyChild Solution were used for the conventional HMIS. **Acceptability Criteria:** The acceptable threshold for reporting completeness is receiving at least 75% of reports.
1.2 Timeliness of facility reporting	Percentage of submitted facility monthly reports (previous 1 year) that are received on time (i.e. by the deadline for reporting).	March 2018 to February 2019 monthly reports of all health facilities implementing MyChild Solution including their submission deadlines were used*. Last 12 months before the implementation of MyChild Solution were used for the conventional HMIS. **Acceptability Criteria:** The acceptable threshold for reporting timeliness receiving at least 75% of reports on time.
1.3 Completeness of indicator data	Percentage of data elements that are filled in and non-zero values for Pentavalent vaccine 3.	March 2018 to February 2019 monthly reports of health facilities implementing MyChild Solution were used.* Last 12 months before the implementation of MyChild Solution were used for the conventional HMIS. **Acceptability Criteria:** The recommended threshold is of at least 75%.
Dimension 2: Internal consistency of reported data		
2.1 Outliers	Number and percentage of subnational units in which one or more of the monthly subnational unit values over the course of one year is a moderate (±2–3 standard deviations from the mean) or extreme (> ±3 standard deviations from the mean) outlier.	MyChild Solution's de-identified electronic register data March 2018 to February 2019 was used. Four health facilities were included in assessing this metric because the rest had used the solution for less than 12 months. The last 12 months before the implementation of MyChild Solution were used for the conventional HMIS. **Acceptability Criteria:** The thresholds for moderate and extreme outliers are ± 2–3 SD from the mean and > ±3 SD from the mean respectively.
2.2 Consistency between related indicators	Pentavalent vaccine dropout rate: a) for the conventional HMIS, this was calculated using the aggregate number of children receiving each vaccination as follows: (Pentavalent vaccine 1 – Pentavalent vaccine 3) / Pentavalent vaccine 1b) for MyChild Solution, this was calculated using the actual number of eligible children based on individual records, i.e. those who received Pentavalent vaccine 1 and are due for Pentavalent vaccine 3 but did not receive it, divided by the total number of children who received Pentavalent vaccine 1.	The dropout was calculated using March 2018 to February 2019 electronic register data for all health facilities. Dropout rates calculated using aggregate doses administered in 2019 (January – December) were used for the conventional HMIS. **Acceptability Criteria:** The dropout values should not be negative.
2.3.1 Verification of reporting consistency through facility survey	Agreement between verified counts for selected indicators in sampled facility records and reported values for the same facilities. Our selected indicator was Pentavalent vaccine 3.	Tally sheets, de-identified electronic register, monthly returns, and DHIS 2 values for December 2018, January 2019, and February 2019 of six randomly selected health facilities were used. Verification factors were calculated. Verification factors are defined as the ratio between reported and verified values (two data sources). A verification factor of 1.0 indicates perfect consistency, less than 1.0 indicates over-reporting, and more than 1.0 indicates under-reporting. Equivalent data sources and same months before the implementation of MyChild Solution were used for the conventional HMIS. **Acceptability criteria:** The acceptable range for verification factors is 0.9–1.1.
2.3.2 Consistency between vaccines recorded in MyChild Health Records Form and electronic records	The percentage of children with vaccine doses recorded on Smart Paper Forms correctly recognized by the system.	Forty-two documents containing 1,527 visits were randomly selected using all document numbers in the system from March 2018 to February 2019. Digital images of SPT forms and digital images of digitized data were compared. This metric is not part of the DQR Toolkit. This metric is not applicable in the conventional HMIS. **Acceptability criteria:** An acceptability threshold of 90% was used.
Dimension 3: External consistency		
3.1. Pentavalent vaccine 3 coverage in comparison with survey coverage	Comparison of pentavalent vaccine 3 coverage from MyChild Solution and conventional HMIS data to national survey results.	MyChild Solution's Pentavalent vaccine 3 coverage in each of the two regions as of the end of February 2019 was estimated and compared with those of the conventional HMIS (2019), and The Gambia Multiple Indicator Cluster Survey (MICS) 2018. **Acceptability criteria:** Values should be within 33% of The Gambia MICS 2018.

* Some health facilities started using MyChild Solution after March 2018. For such health facilities, the number of monthly reports that could be included was less than 12.
NB: The number after a vaccine name is referring to the dose in a series e.g Pentavalent vaccine 3 is referring to the third dose of the Pentavalent vaccine.

2.4. Ethical considerations

This evaluation was a component of a wider project evaluation in The Gambia. All data used in the evaluation is secondary routine service delivery data that was obtained after presenting a detailed data needs sheet that documented the metrics to be evaluated and the specific data required to the data providers. A de-identified version of the original immunisation register was used to carry out analyses requiring individual-level data. Given all data was provided to the evaluation team without the presence of individual identifiers, this evaluation did not require a formal ethics approval.

3. Results

3.1. Dimension 1: Completeness and timeliness

One hundred and forty-six (146) reports were expected from MyChild Solution and 228 reports from the conventional HMIS. All (100%) expected reports were submitted and on time for both MyChild Solution and the conventional HMIS system. The agreed deadline for timely submission of Monthly Returns between the Ministry of Health and Shifo Foundation is the 5th of the subsequent month. Pentavalent vaccine 3 data was also 100% complete in both systems. All fixed and outreach immunisation sessions (n = 510) conducted in April 2018, December 2018, and February 2019 were captured in MyChild Solution resulting in 100% completeness of captured immunisation sessions.

3.2. Dimension 2: Internal consistency

Neither of the two systems had a health facility with moderate or extreme outliers as shown in Table 2. SPT forms and electronic records were more than 99% consistent. The SPT forms and electronic records consistency could not be measured for the conventional HMIS because the system does not use SPT forms and individual electronic records. Pentavalent vaccine 1 – Pentavalent vaccine 3 dropout rate was 3.5% in the conventional HMIS, and 15.8% in MyChild Solution. Dropout rates for both systems are in the expected direction (i.e. not negative). However, MyChild Solution's dropout rate is higher than that of the conventional HMIS. The main reason is that the two systems use different methods for calculating the dropout. MyChild Solution uses the actual number of children who received Pentavalent vaccine 1 and are due for Pentavalent vaccine 3 but did not receive it as the numerator, whilst the HMIS uses the difference between the total Pentavalent vaccine 1 and the total Pentavalent vaccine 3 doses as a numerator.

In MyChild Solution, tally sheets and Monthly Return forms are perfectly (100%) consistent for the selected months in the selected health facilities as shown by the Verification Factors (VFs) in Table 3. In health facilities B, C, and D, there were fewer total doses in the immunisation registers than in the Monthly Returns, though this difference is within the acceptable threshold. There was serious under-reporting from the Monthly Return forms to DHIS-2 for facility A (VF = 1.6) in December 2018, and for facilities A (VF = 1.21), C (VF = 1.32), E (VF – 4.62), and F (VF = 2.5) in January 2019. The analysed months were prior to the DHIS-2 integration of MyChild Solution through the routine use of the DHIS-2 App. Given the recent full routine integration of the MyChild Solution DHIS-2 App, such under- or over-reporting should be zero for MyChild Solution as the monthly vaccine dose totals should automatically be transferred to DHIS-2 from MyChild Solution.

Table 4 summarises Verification Factors of total Pentavalent vaccine 3 doses for the conventional HMIS. Although most (33/54) of the verification factors are within the DQR Toolkit's acceptable limits, a significant number of them (21/54) are outside the acceptable thresholds. Inconsistency between immunisation registers and Monthly Returns had the highest occurrence (16/54), followed by missingness (NA) in one of the data sources (4/54), and inconsistencies between Monthly Returns and DHIS-2 values had the lowest occurrence (1/54). Immunisation registers were not consistently updated when the selected health facilities were using the conventional HMIS.

Table 2
Presence of outliers, related indicators consistency, and electronic records and SPT forms consistency for the conventional HMIS and MyChild Solution.

Metric	Conventional HMIS	MyChild solution
Number and percentage of health facilities with moderate outliers n (%)	0 (0%)	0 (0%)
Number and proportion health facilities with extreme outliers n (%)	0 (0%)	0 (0%)
Penta 1 - Penta 3 dropout rate	4%	16%
SPT forms and electronic records consistency	Not Applicable	99.98%

3.3. Dimension 3: external consistency

Fig. 2 shows Pentavalent vaccine 3 coverage from three different sources for West Coast Region 1 and West Coast Region 2 respectively: The Gambia Multiple Indicator Cluster Survey 2018 (90.4% and 94.9%) [13], the conventional HMIS (92.0% and 94.0%), and MyChild Solution (66.8% and 65.9%). All the coverages are within 33 percentage points of each other, thus within the acceptable limits set by the DQR Toolkit.

3.4. Data quality indicators beyond the DQR toolkit

Table 5 summarises incidences of recording errors for MyChild Solution. The proportion of children recorded as having received more than one dose of BCG in different visits appears to be higher in fixed (1.6%) than outreach (0.8%) immunisation sessions ($P < 0.001$). Please note that this statistic is specific to BCG, so the proportion of children recorded as having received multiple doses of any vaccine is expected to be higher. However, the proportion of children recorded as having received multiple doses of the same antigen in one visit is the same (0.4%) for both types of sessions.

Table 6 summarizes incidence of incomplete indicator level data. The most common form of incomplete indicator level data is not recording the sex of a child (5.93%). Birth dates in the future (1.41%), defined as birth dates being later than clinic dates, is the most frequent form of implausible dates. Please note that this metric may result from both incorrect birth dates and incorrect clinic dates.

4. Discussion

We assessed the quality of routine immunisation data generated using MyChild Solution, an electronic immunisation information system based on SPT, and compared it with the conventional HMIS in place before its introduction. Data quality indicators evaluated using the WHO DQR Toolkit were mostly found to be within the recommended routine immunisation data quality thresholds for both immunisation information systems [9]. MyChild Solution generates more internally consistent data across data sources than the conventional HMIS and has the capability to eliminate inconsistencies between Monthly Returns and DHIS-2 values because of its ability to be integrated with DHIS-2. Most of the inconsistencies in the conventional HMIS are due to not updating immunisation registers, followed by missingness in a data source.

Table 3
Verification factors of total Penta 3 doses between different immunisation data sources using MyChild Solution.

Month/comparison	Verification factors by health facility					
	Health facility A	Health facility B	Health facility C	Health facility D	Health facility E	Health facility F
December 2018						
Tally sheets and monthly returns	1	1	1	1	1	1
Register and monthly return*	1	0.99	0.99	0.93	1	1
Monthly return and DHIS-2	1.6	1.04	0.98	1	1	1
January 2019						
Tally sheets and monthly returns	1	1	1	1	1	1
Register and monthly return*	1	0.97	0.97	0.93	1	1
Monthly return and DHIS-2	1.21	1	1.32	1.01	4.62	2.5
February 2019						
Tally sheets and monthly returns	1	1	1	1	1	1
Register and monthly return*	1	0.97	0.98	0.97	1	1
Monthly return and DHIS-2	1	1	1	1.03	1.01	1

*Register and Monthly Return totals now have a verification factor of 1.0 (100% consistent) across all health facilities and months. The acceptable difference between the two data sources was caused by a technical issue that has been detected and corrected.

Table 4
Verification factors of total Penta 3 doses between different immunisation data sources using the conventional HMIS. NA means data was not available for at least one of the data sources used for the comparison.

Month/comparison	Verification factors by health facility					
	Health facility A	Health facility B	Health facility C	Health facility D	Health facility E	Health facility F
December before the implementation of MyChild Solution						
Tally sheets and monthly returns	1.00	1.00	0.98	1.00	1.05	NA
Register and monthly return	0.03	0.02	0.08	0.90	0.28	0.87
Monthly return and DHS 2	1.00	1.01	1.00	1.26	1.00	NA
January before the implementation of MyChild Solution						
Tally sheets and monthly returns	1.00	1.01	1.01	1.00	0.91	1
Register and monthly return	0.02	0.05	0.15	0.96	0.27	0.72
Monthly return and DHS 2	1.00	1.00	1.00	1.00	1.00	NA
February before the implementation of MyChild Solution						
Tally sheets and monthly returns	1.00	0.97	1.00	1.00	0.97	1
Register and monthly return	0.08	0.00	0.15	0.82	0.35	0.84
Monthly return and DHS 2	1.00	1.00	0.90	1.00	1.00	NA

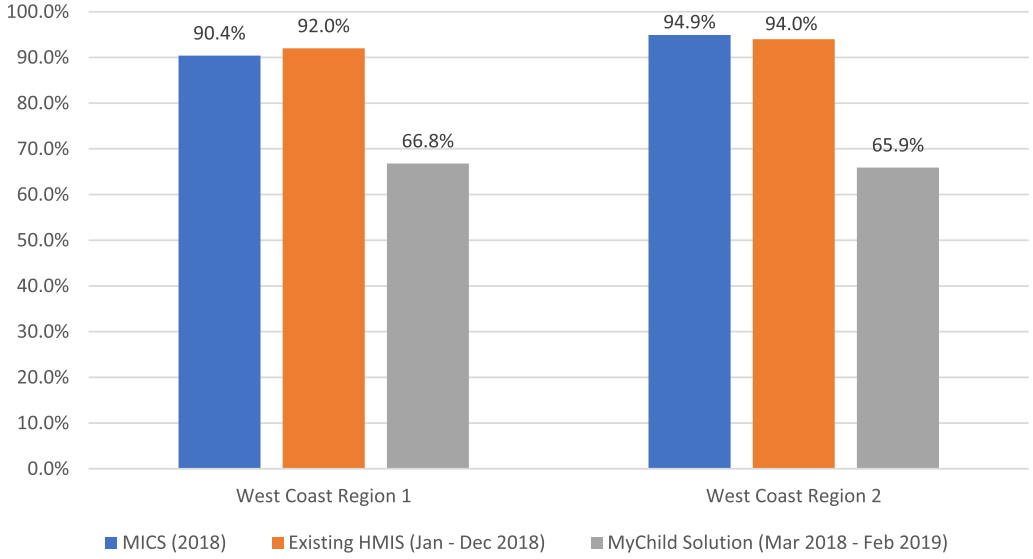

Fig. 2. Third dose of Pentavalent vaccine coverage consistency using different data sources (The Gambia Multiple Indicator Cluster Survey 2018, the existing HMIS and MyChild Solution) in West Coast Region 1 and West Coast Region 2.

Additional metrics evaluated in MyChild Solution showed very encouraging data quality results.

Completeness and timeliness of reporting are commonly evaluated attributes, being cornerstones for data availability. Our findings showed that both MyChild Solution and the conventional HMIS produce perfectly complete and timely data. Peer reviewed SPT data quality assessments are scarce. However, other data quality assessments based on different systems provide helpful comparisons for completeness and timeliness of reporting results. For example, a study conducted in Iran showed that both data completeness and timeliness by the paper-based system were well above 75% [14], the minimum threshold recommended by the

Table 5
Incidences of recording errors in MyChild Solution.

Incidences of recording errors (type of recording error)	Fixed n (%)	Outreach n (%)	Total n (%)
Proportion of children recorded as having received multiple doses of BCG during different visits.	248 (1.6%)	31 (0.8%)	279 (1.4%)
Proportion of children recorded as having received multiple doses of Oral Polio, Pentavalent, Rotavirus, Pneumococcal conjugate, and Measles vaccines during the same visit.	360 (0.4%)	115 (0.4%)	475 (0.4%)

Table 6
Incidence of incomplete data entry and implausible dates.

Type (Population = 65261)	Incidence n (%)
No sex recorded	3781 (5.93%)
No actual birth date recorded	684 (1.05%)
Date of birth in the future*	922 (1.41%)
Children more than 5 years at registration**	115 (0.18%)

* Birth dates were considered to be in the future if they post-dated the clinic date. This could reflect both an inaccurate recording of the birth date or inaccurate recording of the clinic date.
** Children might be registered over five if receiving vit A or deworming services.

WHO DQR Toolkit. A case study that examined the quality of routine data generated by a paper-based system for monitoring priority maternal and newborn indicators available through DHIS-2 in Gombe state, Nigeria found reporting completeness to have on average met the required completeness standard indicated in the DQR Toolkit but not the standard for timeliness [10]. Whereas both the systems evaluated in our study were found to have complete indicator data, indicator data completeness was below the recommended standard for all the immunisation related indicators assessed by Bhattacharya *et al.* (2019) in Gombe state, Nigeria [10]. In addition to the DQR recommended indicators, this evaluation also assessed the completeness of immunisation sessions captured by MyChild Solution. This indicator is automatically tracked by the MyChild Solution system for each health facility. We found that all immunisation sessions were correctly captured in the system. This feature of MyChild Solution gives it an edge over the conventional HMIS because it allows remotely comparing sessions captured by the system against those scheduled for each health facility using the solution's interactive dashboard. This enables quasi-real-time (as forms get scanned) monitoring and follow-up with health facilities from the regional and national levels.

The absence of outliers in both MyChild Solution and the conventional HMIS over the time periods evaluated suggests that there is not a high variation in monthly aggregate data for both systems. In their study conducted in Nigeria, Bhattacharya et al. (2019) reported that more than half of the data elements had outliers in the reference year [10]. In terms of verifying the expected relationship between related indicators, the expected dropout direction was observed in both evaluated systems. In the study by Bhattacharya et al. (2019), the expected relationship between related indicators was not consistently observed [10]. However, it is worth noting that while Bhattacharya et al (2019) measured dropout and coverage like the conventional HMIS, MyChild solution calculates these indicators in a different way". MyChild Solution calculates coverage based on eligible children in the system who have received a given immunisation dose, as opposed to relying on target numbers for the denominator. Calculating vaccination coverage using eligible children in an EIR as a dynamic denominator, as opposed to relying on aggregate estimates, has the potential of providing greater precision in coverage estimates provided that all eligible children are registered and migration is not a significant factor. The existence of health facilities not currently implementing MyChild Solution affects coverage estimates yielded by the system. This is because children registered in the system for earlier doses are captured in the denominator when calculating coverage for later doses, but not in the numerator if they received later vaccine doses in health facilities not implementing MyChild Solution. In a similar logic, non-implementing health facilities influence dropout indicators in MyChild Solution given data is missing for children that receive vaccinations without these being recorded in the MyChild Solution system. This results in an overestimation of the number of defaulters and the dropout rate. This challenge will be overcome if and when the solution is scaled nationally.

For MyChild Solution, we additionally assessed congruence between scanned SPT forms and digitised data in the system as a measure of the accuracy of the digitisation of scanned SPT forms. We found the consistency between the two to be near perfect and well within the range for consistency between different data sources recommended by the WHO DQR Toolkit. According to a similar study, the MyChild Solution system in Uganda correctly processed 97% of vouchers [15]. The data quality metric with the largest disparity between MyChild Solution and the conventional HMIS is the consistency between different routine immunisation data sources. Monthly returns, EIRs, and tally sheets are perfectly consistent in MyChild Solution as they are generated from the same data source. On the other hand, monthly aggregated immunisation values obtained from immunisation registers were mostly much lower than those in the monthly returns in the conventional HMIS. This inconsistency can be explained by health workers not regularly and consistently updating immunisation registers, particularly during busy clinic sessions. A study conducted in Nigeria highlighted substantial discrepancies between paper-based immunisation registers and monthly summaries [16]. Although different studies reported different results, it is evident that the issue of under- and over-reporting is widespread [10,14,16–18]. Failing to record the immunisation history of children in immunisation registers would have implications for retrieving immunisation histories of children when they lose their home-based records or during assessments that rely on immunisation registers at health facilities. It is therefore vital to encourage routine updating of immunisation registers in areas where conventional paper-based health information systems are in place. The under-reporting of data in DHIS-2 is presumably related to manual data entry error as the DHIS-2 MyChild App was developed but not consistently used at the time of the evaluation. This type of inconsistency should be completely eliminated now as MyChild Solution's integration with the DHIS-2 is fully operational and data is therefore directly fed into the DHIS 2 platform. This should have eliminated manual data entry processes which can result in inconsistencies.

While the consistency of data from different data sources are within the acceptable threshold indicated by the DQR Toolkit, differences in coverage values can be explained by the different methods used to obtain these estimates. The Gambia Multiple Indicator Cluster Survey measured coverage through household surveys that rely on home-based records inspections or caregivers' recall if the card was not available or the vaccination was not recorded as given [13]. The conventional HMIS measures coverage using the aggregate number of doses given as the numerator and immunisation targets based on demographic statistics as the denominator. Target-based denominators can be controversial, particularly those set based on sub-national catchment areas. Migrating populations may result in over- or under-estimation of targets [19]. Data quality issues related to the numerator, such as over or under-recording, can also affect coverage. Further, large hospitals with high birth rates and conveniently positioned clinics (e.g. clinics near large markets) will offset coverages, particularly in urban

areas served by several clinics. MyChild Solution, on the other hand, calculates coverage based on individual-level data by calculating the fraction of registered children who were due to receive a specific vaccine and actually received that vaccine. Not being implemented at the national level at this stage, however, artificially lowers coverages and raises dropout rates calculated by MyChild Solution due to the aforementioned issue of "fake defaulters" resulting from non-participating clinics. Scaling MyChild Solution nationally would ensure coverage values are accurately representing all children in the system, while ensuring dropout and coverage rates are based on individual-level data as opposed to targets.

High-quality health data is paramount to health systems planning, implementing new programs, and evaluating their effectiveness [20]. Based on the importance of routinely monitoring data quality of information systems, a number of tools besides the WHO DQR Toolkit have been developed including the Performance of Routine Information System Management (PRISM) tool developed by Aqil et al. (2009) for assessing data quality [21], which has been applied by Hotchkiss et al. (2010) [22], among others. MyChild Solution's reliance on a lower number of data collection tools at the point of care and automated digitisation, as opposed to manual entry processes, results in inherently higher data quality. While it produces high-quality individual-level data, its reliance on paper as opposed to hardware tools at the point-of-care is advantageous in low-resource settings given the frequency of power and internet outages in these settings. This also implies the solution would likely have lower operational costs than most digital solutions that require the use of computers, tablets or smartphones for data collection at service delivery points.

5. Limitations

This evaluation demonstrated the potential for MyChild Solution to improve immunisation data in the Gambia by enabling the availability of individual-level data as well as improving the overall quality of data compared to the conventional paper-based HMIS. One of the key limitations of this evaluation is that the last dimension of the DQR Toolkit, external comparisons of population data, was not assessed. This dimension of data quality evaluates congruence between two population data sources. The dimension has three metrics, none of which were measured in this assessment. The *consistency of population projections* and the *consistency of denominator between program data and official government population statistics* could not be measured as official government population statistics and UN population projections are not estimated by health center catchment area and there is no region/local government area in which all health facilities use MyChild Solution. Therefore, these metrics would be best measured if and when MyChild Solution is implemented in at least all health facilities offering immunisation services in a region/local government area. Further, estimating it at least at the regional level would minimise the possibility of biasing estimates due to people living in one geographical area (for example a given district) and receiving immunisation services in another. The *consistency of population trends* requires three years of data. As MyChild Solution had not been implemented in The Gambia for three years at the time of the evaluation, this metric should be revisited in the future. One metric that could not be included in our assessment due to the sensitive nature of the data required is data accuracy. This would require comparing children's vaccination information in the EIR with the vaccination history of children recorded in home-based records, or to observed administered vaccines, to ensure that SPT forms accurately capture vaccinations being administered.

6. Conclusions

Our assessment has shown that, MyChild Solution, a SPT-based EIR suitable for low-resource settings, produces high quality data with high completeness, timeliness, and consistency compared to the conventional HMIS as measured using the WHO DQR Toolkit. Full integration (completed after data for the evaluation had been collected) of DHIS-2 with MyChild Solution should completely have eliminated inconsistencies between DHIS-2 values and Monthly Return forms as the immunisation data should be automatically fed into the DHIS-2 platform. The solution has the capacity to improve routine immunisation data quality, especially when rolled out nationally. The more internally consistent data produced through MyChild Solution compared to the conventional HMIS demonstrates its potential for supporting data-driven decision-making in immunisation. National scale-up would also enable evaluating external consistency of data produced using the solution. Although the incidence of data quality issues is low, it is important to maintain robust data quality assurance processes including continuously providing feedback to health workers on recording-related issues, monitoring data verification processes, and continuously assessing system-wide data quality metrics to minimise the risk of potential data quality issues.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: A.S. and M.I.G. were part of an external evaluation team contracted to evaluate different aspects of MyChild Solution including data quality in 2019.

Acknowledgments

We wish to thank staff of The Gambia Expanded Program on Immunisation, Shifo Foundation Sweden and its funders and implementation partners for their support, particularly Sanni Kujala, Shahnnoza Eshonkhojaeva, Vasily Loginov, Yorro Cham, Sarjo Gassama and Dawda Sowe for their support and assistance in conducting this evaluation. We would also like to acknowledge Robert Ninsion, who led two other components of the full evaluation, and Samba Bah, Alieu Jallow, Ismaila Kasseh, and Ousman Bah, who assisted in the collection of data.

Author contributions

Conceptualisation, A.S. and M.I.G.; Methodology, A.S. and M.I.G.; Validation, A.S. and M.I.G.; Formal Analysis, A.S with input from M. I.G.; Writing-Original Draft Preparation and Reviewing, A.S. and M. I.G.; All authors approved the final version of the manuscript.

Funding

The evaluation of MyChild Solution in The Gambia, of which the data quality evaluation that culminated into this publication was a part of, was funded by MyChild Project (in The Gambia) through Shifo Foundation, Sweden.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.vaccine.2020.07.074.

References

[1] World Health Organization, editor. Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies. Geneva: World Health Organization; 2010

[2] Friedman DJ, Parrish RG, Ross DA. Electronic health records and US Public health: current realities and future promise. Am J Public Health 2013;103:1560–7. https://doi.org/10.2105/AJPH.2013.301220.

[3] Immunization data: evidence for action. Realist review of what works to improve data use for immunization, evidence from low- and middle-income countries. Seattle: PATH; Washington, DC: PAHO; 2019.

[4] Harrison K, Rahimi N, Carolina D-H. Factors limiting data quality in the expanded programme on immunization in low and middle-income countries: A scoping review. Vaccine 2020;38:4652–63. https://doi.org/10.1016/j.vaccine.2020.02.091.

[5] Trumbo SP, Contreras M, García AGF, Díaz FAE, Gómez M, Carrión V, et al. Improving immunization data quality in Peru and Mexico: Two case studies highlighting challenges and lessons learned. Vaccine 2018;36:7674–81. https://doi.org/10.1016/j.vaccine.2018.10.083.

[6] Danovaro-Holliday MC, Ortiz C, Cochi S, Ruiz-Matus C. Electronic immunization registries in Latin America: progress and lessons learned. Rev Panam Salud Publica Pan Am J Public Health 2014;35:453–7.

[7] Seymour D, Werner L, Mwansa FD, Bulula N, Mwanyika H, Dube M, et al. Electronic immunization registries in Tanzania and Zambia: shaping a minimum viable product for scaled solutions 7. Front Public Health 2019. https://doi.org/10.3389/fpubh.2019.00218.

[8] Pan American Health Organization. Electronic immunization registry: practical considerations for planning, development, implementation and evaluation. Washington, D.C.: PAHO; 2017

[9] World Health Organization. Data quality review: a toolkit for facility data quality assessment: module 1: framework and metrics. Geneva: 2017

[10] Bhattacharya AA, Umar N, Audu A, Felix H, Allen E, Schellenberg JRM, et al. Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: A case study from Gombe State, Nigeria e0211265. PLOS ONE 2019;14. https://doi.org/10.1371/journal.pone.0211265.

[11] StataCorp. Stata Statistical Software. College Station, TX: StataCorp LP; 2017.

[12] R Core Team. R: A language and environment for statistical computing. Vienna, Austria; 2019.

[13] The Gambia Bureau of Statistics (GBOS). The Gambia multiple indicator cluster survey report, 2018. Banjul, The Gambia: Gambia Bureau of Statistics (GBOS); 2019.

[14] Karami M, Khazaei S, Babaei A, Yaghini FA, Gouya MM, Zahraei SM. Accuracy and quality of immunization data in Iran: findings from data quality self-assessment survey in 2017. BMC Health Serv Res 2019;19:371. https://doi.org/10.1186/s12913-019-4188-9.

[15] Äijö A, Schäffner I, Waiswa P, Kananura RM, Tessma MK, Hanson C. Assessment of a novel scanner-supported system for processing of child health and immunization data in Uganda 20. BMC Health Serv Res 2020. https://doi.org/10.1186/s12913-020-05242-1.

[16] Omoleke SA, Tadesse MG. A pilot study of routine immunization data quality in Bunza Local Government area: causes and possible remedies. Pan Afr Med J 2017;27.

[17] Hu Y, Zhang X, Li Q, Chen Y. Auditing the immunization data quality from routine reports in Shangyu District, East China 13. Int J Environ Res Public Health 2016. https://doi.org/10.3390/ijerph13111158.

[18] Nsubuga F, Luzze H, Ampeire I, Kasasa S, Toliva OB, Riolexus AA. Factors that affect immunization data quality in Kabarole District 13. Uganda. PLoS ONE 2018. https://doi.org/10.1371/journal.pone.0203747.

[19] Nguyen TD, Dang AD, Damme PV, Nguyen CV, Duong HT, Goossens H, et al. Coverage of the expanded program on immunization in Vietnam: Results from 2 cluster surveys and routine reports. Hum Vaccines Immunother 2015;11:1526–33. https://doi.org/10.1080/21645515.2015.1032487.

[20] Murray CJL, Lopez AD, Wibulpolprasert S. Monitoring global health: time for new solutions. BMJ 2004;329:1096–100. https://doi.org/10.1136/bmj.329.7474.1096.

[21] Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan 2009;24:217–28. https://doi.org/10.1093/heapol/czp010.

[22] Hotchkiss DR, Aqil A, Lippeveld T, Mukooyo E. Evaluation of the Performance of Routine Information System Management (PRISM) framework: evidence from Uganda. BMC Health Serv Res 2010;10:188. https://doi.org/10.1186/1472-6963-10-188.