Machine Learning for Predicting Default from HIV Services in Mozambique

OpCon Mozambique Final Report

PEPFAR

HRSA Health Resources & Services Administration

dimagi

ICAP GLOBAL. HEALTH. ACTION. Columbia University Mailman School of Public Health

This project is supported by the U.S. President's Emergency Plan for AIDS Relief, through the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) under grant UJ7HA31180, Optimizing Momentum toward Sustainable Epidemic Control. This information or content and conclusions should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS or the U.S. Government.

Table of Contents

Executive Summary

Background

Successful HIV programs must effectively identify people living with HIV, link them to high quality treatment, and retain them in lifelong care. Patient education, psychosocial support, adherence counseling, peer support, and patient-centered models of care are important interventions to maximize retention; these are generally offered to all recipients of care, rather than to those at high risk of missing appointments and/or being lost to follow-up. In this era of limited resources, strategies are needed to easily and accurately target the subset of individuals at highest risk for defaulting with tailored interventions, thus enhancing programmatic effectiveness and efficiency.

ICAP at Columbia University (ICAP) has worked in Mozambique since 2004, and now supports a concerted government effort spanning all levels of the health system to address the dual health threats of HIV and tuberculosis (TB). Service provision, continuous quality improvement, and development of health facility and laboratory infrastructure form the core of these efforts, with innovative, targeted approaches driving progress in human resources for health, optimized antiretroviral (ARV) regimens, and utilization of new technology to improve patient retention in care.

In partnership with the Mozambique Ministry of Health (Ministério da Saúde), the U.S. Health Services and Resources Administration (HRSA), and the U.S. Centers for Disease Control and Prevention (CDC), ICAP worked with Dimagi to leverage the existence of an existing mobile health application called CommCare to test the use of a predictive algorithm to identify patients at higher risk of default. CommCare, and its Mozambique-specific version, *Infómovel,* was developed by Dimagi to support community-based HIV and TB services in Mozambique. The main front-line users of the app are facility-based case managers (*gestor de casos*) and community health workers (*activistas*).

The OpCon Mozambique project was designed to apply Machine Learning (ML) to predict the risk that individuals receiving HIV treatment will "default" - e.g., fail to return for a scheduled appointment - using Infómovel. Key project activities included: development of the ML algorithm; preliminary design of a user interface (UI) or "flag" within Infómovel; and rapid field testing to elicit user feedback. The project had three main activities:

- A scoping visit in May 2018 to explore Infómovel utilization, activista workflow, and data management;
- Programming the ML algorithm and assessing its predictive capabilities by generating a ML dataset with data from ICAP-supported health facilities;
- A field-testing visit in September 2018 to obtain activista feedback on the idea of applying ML to generate an intuitive risk "flag" into the Infómovel user interface.

Scoping Visit Insights

A notable finding from the baseline scoping visit was the fact that Infómovel can only accommodate the introduction of patients newly enrolled on ART who are receiving supportive home visits. As a consequence, health workers use OpenMRS to generate weekly lists of persons who miss appointments (including, but not limited to those in Infómovel). These OpenMRS lists are used to trigger outreach tracing activities, rather than the defaulter follow-up functionality in Infómovel, which is not currently in use. This limited the data

available for the ML analysis. In response, ICAP provided anonymized data on missed appointments from OpenMRS to be included in the analysis.

Machine Learning Results

The ML algorithm was programmed to identify the 20% of patients at highest risk of default. Dimagi then used data from four ICAP-supported health facilities in Nampula and Zambézia provinces, to develop a cleaned dataset containing 1,547 scheduled visits from 491 unique patients. Table 1 compares the actual rate of visits missed by 7, 10, 14 and 28 days with the rate of missed visits for the "high risk" patients flagged by the ML algorithm. In row 1, for example, the actual rate for all patients was 31.8%. In contrast, the rate for patients "flagged" by the ML algorithm was 48.3%.

Table 1

# days by which a visit was missed	Rate of missed appointments for all patients	Rate of missed appointments for patients in ML-predicted "riskiest" 20%
7	31.8%	48.26%
10	28.3%	43.01%
14	26.2%	39.98%
28	19.8%	34.77%

User Feedback Results

The Dimagi team discussed the feasibility and acceptability of the prediction app and its UI with activistas during the field-testing visit. Insights included:

- There is a high level of user acceptance for the existing Infómovel application.
- However, many activistas are not familiar with the current Infómovel workflows for patients who miss appointments, since they use lists generated by OpenMRS, which allow them to identify individuals who miss appointments other than those in Infómovel.

Conclusion

Despite the challenges of data management and utilization within Infomovel and MRS, the ML algorithm was able to identify a subset of patients at higher risk of default, providing proof-of-concept for the project. Users were open to the idea of a 'flag' within Infomovel, and interested in learning more about how it would be used. A larger and cleaner dataset will improve the model, and development of "high-risk" workflows will enable tailored interventions to be provided to patients at high risk of default. These promising results were shared with the Ministry of Health, the U.S. Centers for Disease Control and Prevention (CDC), the U.S. Health Resources and Services Administration (HRSA), and the Dimagi teams implementing Infomovel in Mozambique.

Recommended next steps include:

- Integration of OpenMRS and Infomovel to enable a clearer dataset
- Uploading data on individuals who miss appointments to Infomovel to enable differentiated follow-up workflows

- Development of a Supervisory App to monitor the data
- Monitoring data collection stability to re-run ML predictive models

Section 1: Background and Context

Background and Rationale

Successful HIV programs must effectively identify people living with HIV, link them to high quality treatment, and retain them in lifelong care. Patient education, psychosocial support, adherence counseling, peer support, and patient-centered models of care are important interventions to maximize retention; these are generally offered to all recipients of care, rather than to those at high risk of missing appointments and/or being lost to follow-up. In this era of limited resources, strategies are needed to easily and accurately target the subset of individuals who are at highest risk for defaulting with tailored interventions, thus enhancing programmatic effectiveness and efficiency.

Identifying patients at high risk of default enables HIV programs to provide tailored interventions, but timely identification can be difficult using status quo routine data collection and systems. Furthermore, although broad categorical risk information may be available to frontline health workers – e.g., "newly infected adolescent girls are at high risk of defaulting" – individualized patient-specific data are likely to be more accurate.

Machine learning (ML) provides a unique pathway to analyze multiple factors contributing to individual risk, and has the potential to place decision-making power in the hands of frontline workers in real time. Dimagi has previously used predictive ML technology to generate risk scores for default from HIV treatment, using data points from 6 countries in sub Saharan Africa to predict the risk of an individual patient defaulting. Results from these studies show that 20% of patients can account for 60% of the risk of defaulting, indicating that the use of ML could guide programs and frontline health workers to direct resources to the "riskiest" population.

In this project, Dimagi partnered with ICAP at Columbia University (ICAP), the U.S. Health Resources and Services Administration (HRSA), the U.S. Centers for Disease Control and Prevention (CDC) and the Ministry of Health of Mozambique to assess the feasibility and acceptability of adapting the defaulter tracking risk score to the Mozambique setting, building on Dimagi's existing portfolio of mHealth activities in the country, including the Infómovel mobile health application, and using data that is already being collected by Infómovel and OpenMRS.

The objective of this short project was to explore the feasibility and acceptability of:

- Developing a ML predictive algorithm to identify the subset of patients at highest risk of default;
- Programming a "flag" to identify these high-risk patients in the Infómovel user interface; and
- Obtaining feedback on the initiative from frontline health workers to assess the likelihood that they would be able to use the flag to tailor services to high-risk patients.

The ultimate goal would be to improve resource allocation, allowing for efficient targeting of adherence support for at-risk patients.

ICAP in Mozambique

In 2004, ICAP was awarded funds by the U.S. Government to partner with the Ministério da Saúde (MISAU) to support the scale-up of HIV prevention, care, and treatment services. Initially, ICAP focused on providing national-level technical support to MISAU in policy and protocol development and clinical training as well as supporting a limited number of day hospitals (designated HIV care and treatment facilities). To mitigate the shortage of Mozambican doctors with clinical experience in HIV, ICAP brought in experienced advisors and technical officers to model, mentor, train, and coach their counterparts at MISAU, at the Direcções Provinciais de Saúde (DPS; the provincial health directorates) and the Direcções Distritais de Saúde (DDS; district health directorates) and supported health facilities on the start-up, delivery and management of HIV services. Early on, formal training at the day hospitals and on-site ICAP technical support established a platform from which ICAP could expand training, support, mentorship, and supervision as the number of supported facilities expanded.

Over time, ICAP support evolved to meet the broadening national program's needs and priorities. ICAP now supports health systems strengthening and institutional capacity building as well as comprehensive HIV-related services for maternal and child health, HIV counseling and testing, laboratory, PMTCT, early infant diagnosis, HIV-exposed infant follow-up, first- and second-line adult and pediatric ART, health facility pharmacies, TB infection prevention and control, blood safety, psychosocial and adherence support and positive prevention programs, nutritional support, as well as the integration of TB and cervical cancer screening into HIV services. In addition, ICAP facility support expanded in accordance with MISAU's goal to have one ART facility in every district. Since 2008, ICAP has also been the lead technical partner in designated provinces and districts.

Since 2009, ICAP facility support has contributed to the national decentralization strategy of shifting HIV services from overburdened specialty referral facilities to primary health care clinics that are closer to the patients they serve. ICAP has further expanded facility support to peripheral health centers and has continued to work alongside DPS and DDS to operationalize policy and roll out care and treatment guidelines. At the national level, ICAP participates in technical working groups and supports the development and revision of national guidelines and policies. ICAP also works closely with MISAU to develop, evaluate, and roll out innovations that enhance the quality, effectiveness, and efficiency of HIV prevention, care, and treatment services. In less than a decade, the government of Mozambique has successfully expanded comprehensive HIV prevention, care, and treatment services. ICAP has been a key PEPFAR partner in that success.

Mozambique Context

The OpCon Mozambique project was designed to apply ML to predict the risk that individuals receiving HIV treatment will "default" - e.g., fail to return for a scheduled appointment - using Infómovel, a version of Dimagi's CommCare mobile health application adapted for use in Mozambique. The main front-line users of the app are facility-based case managers (gestor de casos) and community health workers (activistas).

Infómovel is a Commcare application developed by Dimagi to support community-based HIV and TB patient management by frontline workers in Mozambique. The program's main objective is to improve the health and care of HIV and TB patients by improving community-level services such as adherence support, tracking of people who miss appointments, and community-based HIV and TB testing. Dimagi currently works with five clinical partners in Mozambique (ICAP, EGPAF, Ariel, FGH and CCS) to support the

mobile phone-based application, which assists activistas to provide home-based services to PLHIV, including adherence support, tracking of people who miss appointments, PMTCT services and TB screening. Patients are enrolled at the health facility, and then assigned to activistas who do household visits using the mHealth application. If patients miss a clinic visit, this triggers a "follow up" workflow for the activistas.

Figure 1: Infómovel in Mozambique

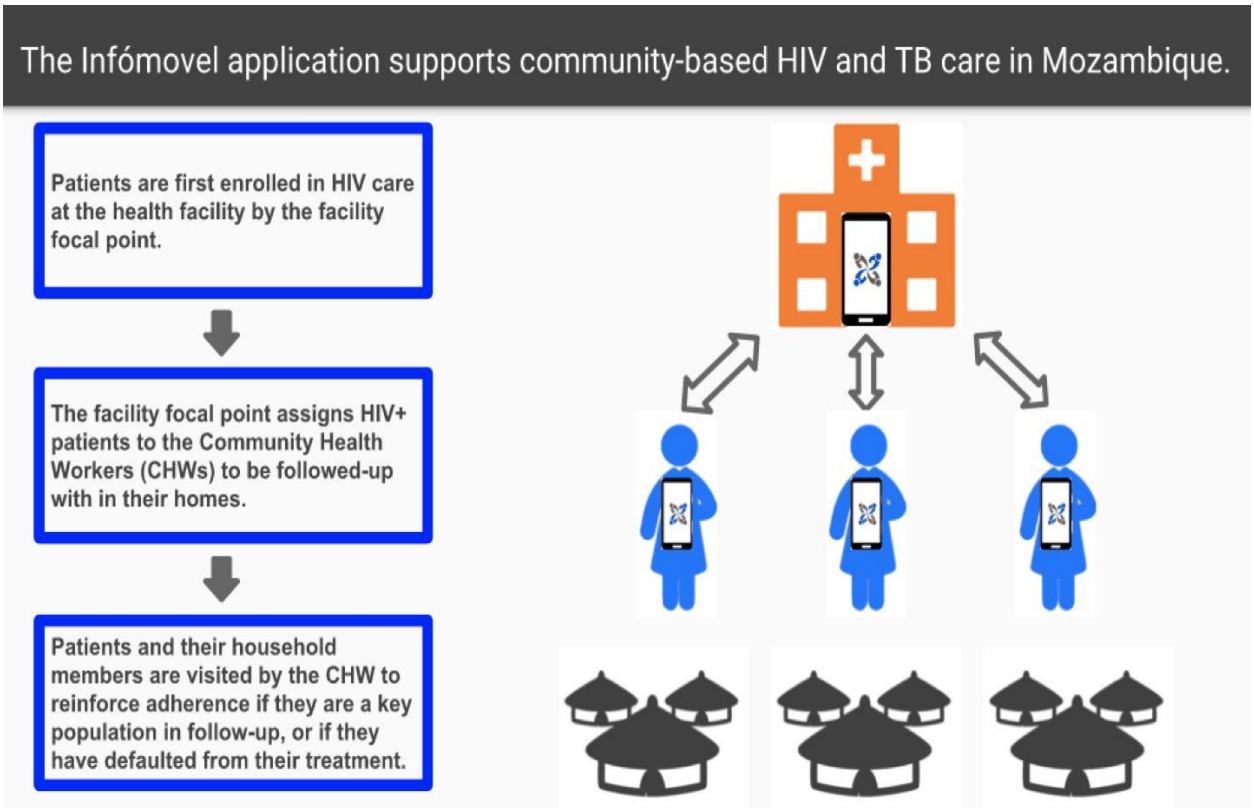


Figure 2: Current tracking workflow for individuals who miss appointments

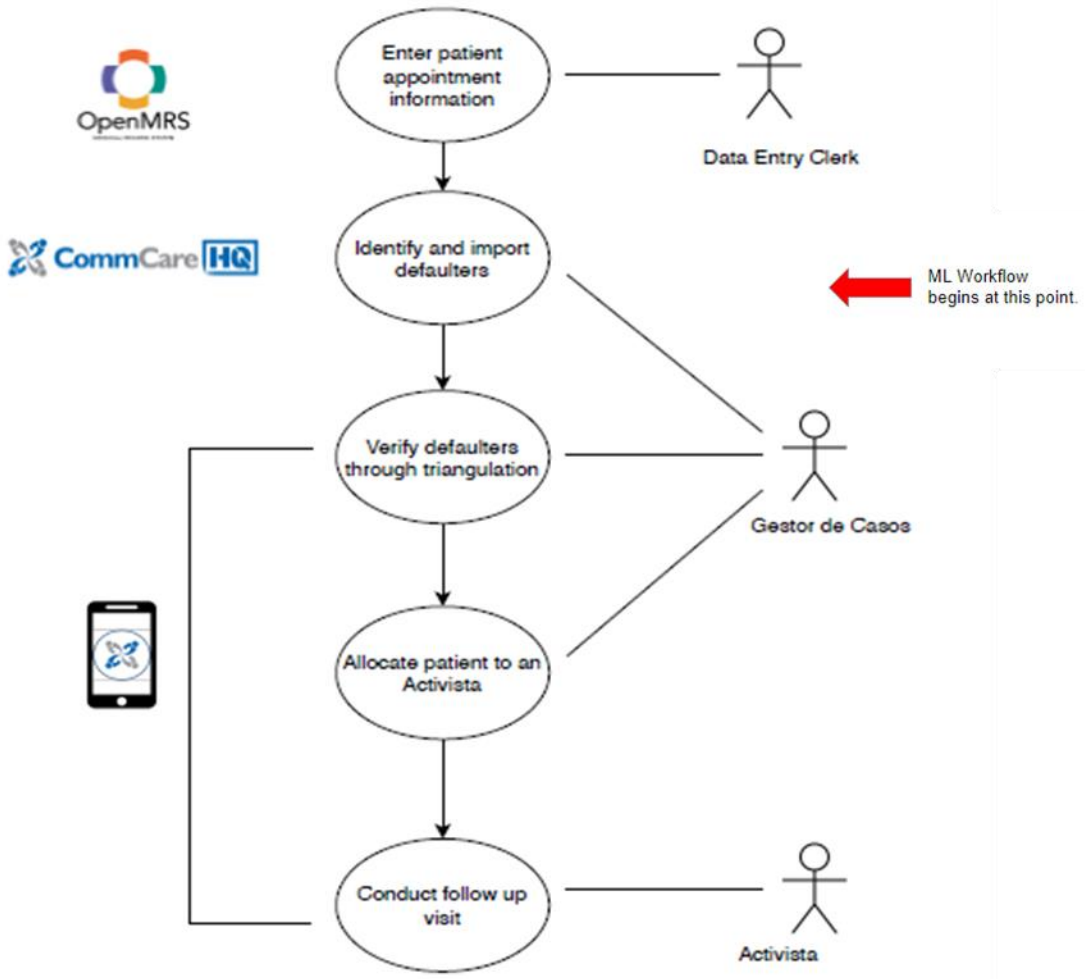


Table 2: ICAP-supported health facilities using Infómovel

Province	Facility Name	Location	# patients on ART	# of CommCare users
Nampula	Muhala Expansão	Nampula city	4152	10
Nampula	Akumi	Nacala Porto district	2949	10
Zambézia	Namacata CS	Nicoadala District	1331	10
Zambézia	CS 16 de Junho	Mocuba District	1664	10

OpenMRS

Health facilities in Mozambique currently use OpenMRS as the Management Information System (MIS) for registering and tracking all patients. This has, for the most part, replaced the earlier ePTS system. OpenMRS captures patients' demographic information, medical history (including laboratory testing) and visit consultation fields, which include the dates of scheduled and actual clinical and medication pick-up visits. Infómovel is used to capture information about home visits and community-based services. A database that merged information from OpenMRS and Infómovel would be an optimal way to capture comprehensive information about missed visits and retention, but this is not yet available.

Section 2: Scoping Visit & Refined Theory of Change

In May 2018, the Dimagi team (Seema Kara and Anushree Banerji) visited Mozambique to explore Infómovel–related workflows, gain insights into related data and patient management, and refine project plans, in consultation with ICAP at MOH.

The goal of the trip was to:

- Understand how data from the Infómovel application travels and is used by different cadres of health workers, and how it interacts with the OpenMRS system;
- Understand how activistas and gestor de casos prioritize patients and assess risk of defaulting;
- Discuss health provider insights on timing of defaulting, reasons for defaulting, and barriers/facilitators to retention in care

The team met with ICAP and MOH in Maputo, with the District MOH (DPS) in Nampula province, and with ICAP staff and health workers at two health facilities in Nampula (Muhala Expansão and Akumi, see Table 2). Details can be found in Appendix 1.

Scoping Visit Takeaways

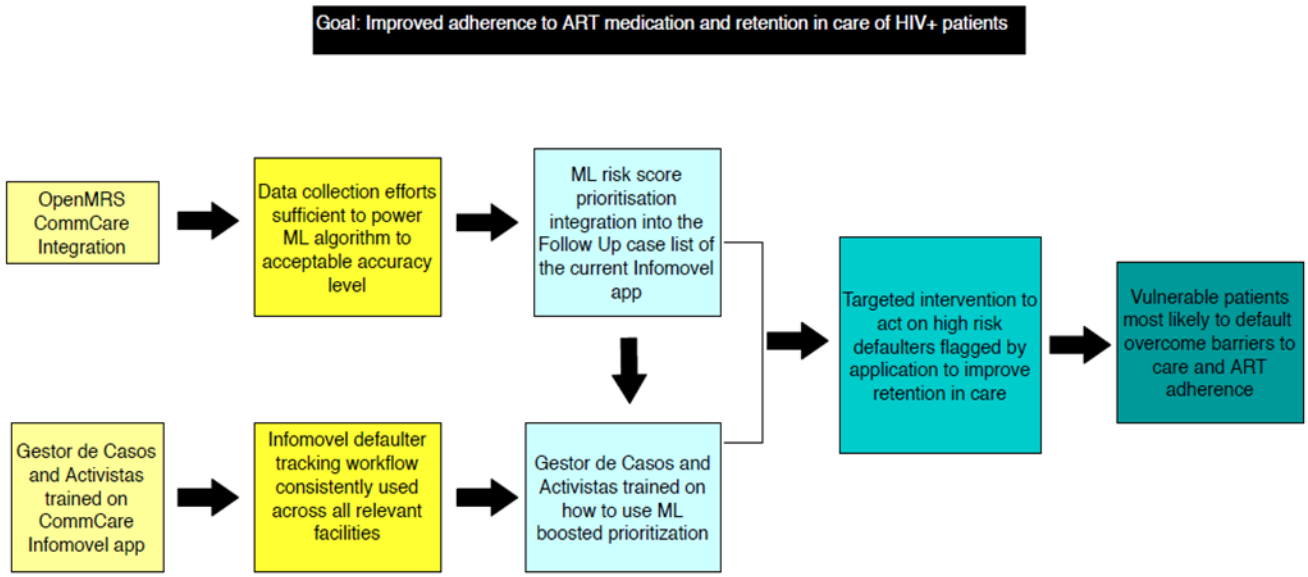
- The Infómovel application was extensively used for tracking active patients, and activistas routinely used the follow-up module to document home visits;
- In contrast, the Infómovel defaulter module ("Busca Activa") was not being used as anticipated and people who miss appointments were identified via OpenMRS rather than Infómovel. This limited usage of the Infómovel Loss to Follow Up (LTFU) forms;
    - NB that sites conducted rigorous tracking and follow-up of people who miss appointments, but the relevant information was not captured in Infómovel, which posed an unexpected challenge for the planned project
- Creation of new patient files on Infómovel by gestor de casos was working well. The timelines for entry varied between facilities but a defined process was being followed at both sites;
- For existing patients, gestor de casos manually generated "appointment due" lists from OpenMRS, triaged with pharmacists and activistas, and then identified the list of patients to prioritize for home visits. Updated dates of consultations and medication pick up were being entered in OpenMRS, not Infómovel.

- There was regular usage of Infómovel data exports and reports for tracking application usage – particularly Patient exports and Home Visit form exports. There was some variation related to which reports or data each facility downloaded or prioritized.

Implications for the OpCon Tracking Project:

The routine use of Infómovel to document home visits and active patient tracking was a promising finding. However, the fact that information about missed visits was being manually generated and was uploaded to Infómovel presented a challenge. The Dimagi team updated its plans, including reaching out to ICAP for access to OpenMRS data for the ML programming activity (see below). Dimagi also finalized a Theory of Change for the project (Figure 3).

Figure 3: Theory of change



Section 3: Machine Learning and Data Analysis

Background: Machine Learning Overview

Machine learning (ML) is a data analytics technique that allows "learning" to take place without being explicitly programmed. As new data enters the system, the ML model is able to adapt and learn. ML algorithms find patterns in data that help users gain insight and make predictions. When applied to complex tasks such as tracing people who miss appointments within a health system, the goal is to use these methods to help understand data generated by a system and present this in a way that can be integrated into specific, relevant tasks.

In this project, the ML models looked at risk scores amongst recipients of HIV treatment who missed appointments, with the goal of assessing the burden of risk across a group of patients, and identifying, for example, the "riskiest" 10-20% of patients, e.g., those who carry 50% or more the risk of defaulting. At the facility level, this can enable providers to flag patients who are at the highest risk of default, and to target them appropriately with additional support.

Machine Learning Activities and Preconditions

In the theory of change, the goal of using ML predictions is connected to both the users of the mobile Infómovel application as well the data system in which Infómovel is embedded. To reach the goal of an end user successfully using a risk score, pre-conditions for both actors/information systems need to be met. These are described in Table 3.

Table 3: Machine learning activities and preconditions

Pre-condition	Assumption
1. Problem Identification	
Predictor (ML target) defined	The predictor is relevant and directly linked to the long-term goal.
Data sources to inform predictor are identified	In a new system, data sources to back up the predictor must be mapped, to support the feasibility of the chosen predictor
End user workflow mapped	Learning from *status quo* practices, we are able to logically fit the predictor into the workflow
2. Create ML unified view dataset	
Obtain necessary data from multiple data sources	Minimum data conditions to create an ML-ready dataset are met by partner organization
Verify accuracy of appointment-based data	We can identify a process to validate key data points, such as appointment data, to ensure we are as "close to the truth" of a system property as possible
3. Design prototype to "operationalize" risk score	
Identify use cases for integrating risk score	There is an established practice/SOP across facilities for missed visit tracking workflow to use to design the ML risk-score.
UAT (User Acceptability Testing) informs usability threshold	A risk score gives the user previously unknown information about their case list. The high-risk patients are easily and intuitively identified with minimal guidance. User feedback validates integration of the ML risk-score into day to day activities and informs design iteration.
Involve multiple stakeholders to define their role within this process	The system comprises of moving parts defined by end users in the system and their distinct contributions based on their activities.

Establish the feedback loop of entering and updating a patient's risk score	The risk score can be updated or refreshed via a combination of automation and manual data entry. Investment in end user training is necessary. Risk of inaccurate results is understood by partner organization if protocols are not followed.

OpCon Machine Learning Objectives

As described above, the objective of the ML process for the OpCon project was to develop a model to assign a defaulter risk score to patients registered on the Infómovel domain of ICAP-supported health facilities. This information would then be used to "flag" these patients in the Infómovel application user interface (UI), to enable health workers to target them appropriately.

In consultation with the project team (ICAP, CDC and HRSA), we narrowed the definition of default to mean "visits that occurred seven days or more after the scheduled date", which is the definition of "Faltoso" in Mozambique's national HIV program. Our main objective was therefore to be able to predict which patients would be at highest risk of missing appointments by seven days. However, to enrich the analysis, we also included data for 10, 14 or 28 day missed appointment cutoffs.

Data Collection and Cleaning

We reviewed anonymized patient data from Infómovel and OpenMRS at the four participating health facilities, which were the ICAP-supported ART clinics using Infómovel in Nampula and Zambézia Provinces (Table 2). Data cleaning processes differed for each export due to their different structures. The steps from the final export, which included July and August 2018 data for the four sites, included:

1. **One row per unique National Identifier (NID):** We first eliminated 8.6% (N=66) of the rows from the OpenMRS export to ensure that there were no duplicate entries for NIDs. We selected the "last" row for any duplicates.
2. **No visit data:** From the resulting data set, we eliminated 7.5% (N=49) rows because of insufficient visit data. This occurred when there were no months with a scheduled visit for a patient.
3. **Untracked NIDs:** Finally, an additional 24% (N=159) of rows were excluded because the NIDs did not match with any first visit forms in Infómovel.

The final dataset contained 1,547 scheduled visits from 491 unique patients (Table 4, Table 5).

Table 4: Number of unique patients with valid data, per site

Site	Patients	Visits
Akumi (Nampula)	313	910
Muhala (Nampula)	146	487
Namacata (Zambézia)	19	88
16 de Junho (Zambézia)	13	62
Total	**491**	**1,547**

Model Inputs: Patient-specific Variables

The information available to the predictive algorithm included the following variables:

- Age
- Sex
- Occupation
- Location of household
- Number of household members
- Number of children
- Household members under the age of 5
- Household members aged 5-14
- Number of partners at home
- Number of children tested
- Patient education level
- Number of pills per day
- Disclosure of HIV status (yes/no)
- Availability of a confidant (yes/no)
- Documented phone number
- Good vs. poor adherence (calculated based on in built logic)
- Patient agreed to visit health facility (Y/N)
- Patient's sexual partners outside the household (Y/N)
- Reasons for not finding patient
- Steps taken to find patient
- Date of last medication pick up
- Scheduled date of medication pick up

Table 5: Illustrative patient demographics

Demographic	Value	Frequency	%
Sex	Male	179	36%
	Female	312	64%
Age	18-20 years	46	9%
	20 - 29 years	175	36%
	30 - 39 years	146	30%
	40 - 49 years	76	15%
	50 years and older	48	10%
Occupation	Domestic	265	58%
	Farmer	80	18%

	Informal	47	10%
	Student	26	6%
	Other	38	8%
HIV status disclosed	Yes	395	80%
	No	93	19%

Machine Learning Methodology

As above, the goal was to use demographic and historical visit data to predict late visits, where a visit is defined as the patient physically attending the facility for clinical services, laboratory services and/or to collect medications, and a late visit is defined as a visit that is seven or more days late. There were a number of different iterations performed, but each followed roughly the same methodology.

1. **Train-test split:** Starting with the full data set of 1,547 scheduled visits, data was split into training and test sets. Borrowing from the standard 10-fold cross-validation strategy, 10% of the initial dataset was set aside for testing the finalized model, with the remainder reserved for training the model.
2. **Class balancing:** The next step was to balance the classes (defaulter and non-defaulter) in the training set. The SMOTE[1] approach was used to achieve a more equal balance.
3. **Training the model:** A number of different approaches were attempted during the exploration phase. The Random Forest technique, a popular ensemble model that uses a 'forest' of randomized decision trees to produce a continuous score, performed best without overfitting.
4. **Testing the model:** The final step was to test the model. This happened differently during the two different phases:
    a. **Exploration phase:** During the exploration phase, 10-fold cross-validation was used to test models internally. The original 10% was only used to test the best model that came out of the internal cycles.
    b. **Finalization phase:** During the final phase, models were checked against the set-aside test set.
5. **Repeat:** There were many cycles where this methodology was repeated. Most notably, during the final phase, each run was repeated 500 times to ensure that the results captured were an accurate representation of the final algorithm and not a product of the randomization split.[2]

Results

The initial analysis explored the results using different definitions of "default" – e.g., 7, 10, 14 and 28 days. In each case, the rate of missed appointments was higher for the subset of patients identified by the ML algorithm as "high risk" than it was for the group as a whole. Table 6 compares the rate of missed appointments for all patients with the rate for those in the ML-predicted riskiest 20% of patients.

---

[1] SMOTE: Synthetic Minority Over-sampling Technique

[2] This is sometimes referred to as Monte Carlo cross-validation. We found 500 cycles to be sufficient to produce a clear representation of the performance of the algorithm

Table 6: Missed appointment rate for all patients vs. "high risk" patients using different definitions of default

# days by which a visit was missed	Rate of missed appointments for all patients	Rate of missed appointments for patients in ML-predicted "riskiest" 20%
7	31.8%	48.26%
10	28.3%	43.01%
14	26.2%	39.98%
28	19.8%	34.77%

Figure 4 shows the results in graph form. The graph is constructed by taking the output of a model, which is a set of risk scores, and ranking them from most to least risky of missing a visit by more than D-number of days. On the graphs in Figures 4a-4d, the axes represent the following:

- The x-value represents the predicted relative risk of default. For example, X = 20%, represents the 20% of patients with the highest predicted risk of missing visits.
- The y-value represents the actual risk of a visit being missed

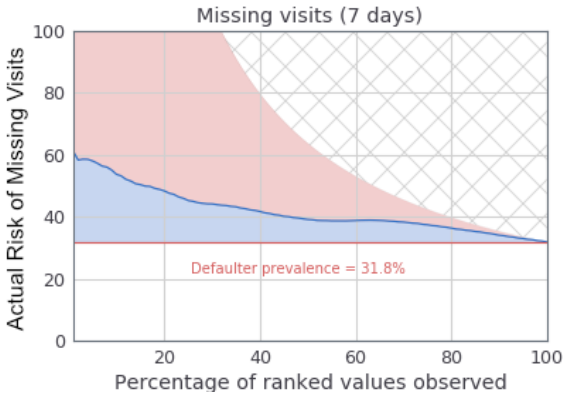
The red line in the graphs represents the average risk of default for all patients[3]. In order to be successful, the ML algorithm must be more precise than this average risk; e.g., should be able to identify the subset of patients at higher-than-average risk.

The blue line in the graphs represents the proportional risk (PR) score, at varying levels of predicted risk. For example, if a point along the blue line is at 20% on the X axis (predicted risk) and 30% on the Y axis (actual risk), the interpretation is that the 20% of patients at highest predicted risk have a 30% chance of missing a scheduled visit. Thus, the blue shaded region represents how well the algorithm performs compared to what we would expect from a random ranking. The red shaded region represents the room for improvement of the algorithm--using the theoretical maximum for any given x-value based on prevalence. The hatched region is unattainable. It is not possible to look at more default values than the prevalence of the full sample.

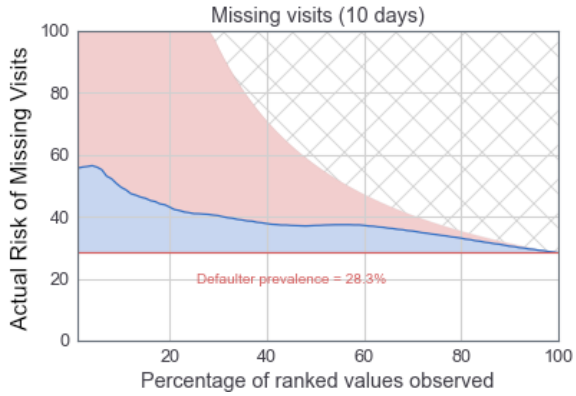
To elaborate, in Figure 4a, we can see that X=20% roughly aligns with Y = 50%. This is interpreted to mean that patients in top 20% risk group had a 50% chance of missing visits by at least 7 days.

Figure 4: (a) Defines a missing visit as 7 days, while (b, c, d) define it as 10, 14 and 28 days, respectively.

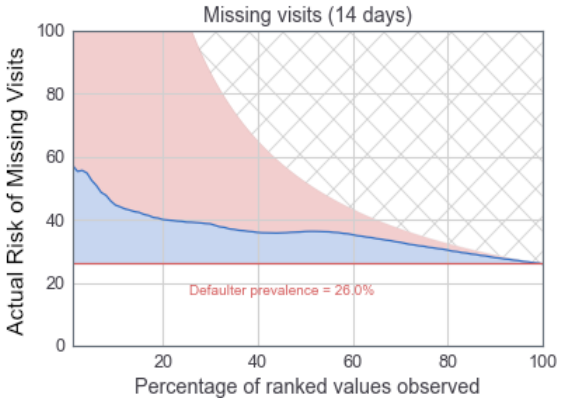
[3] Recall that each graph is the result of Monte Carlo cross-validation with 500 runs, which is why the test set means converge to the prevalence



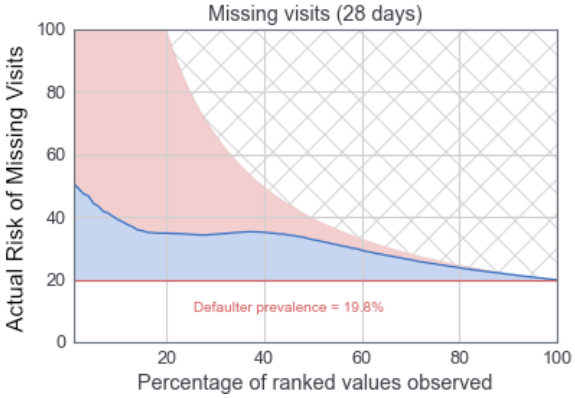
4a



4b



4c



4d

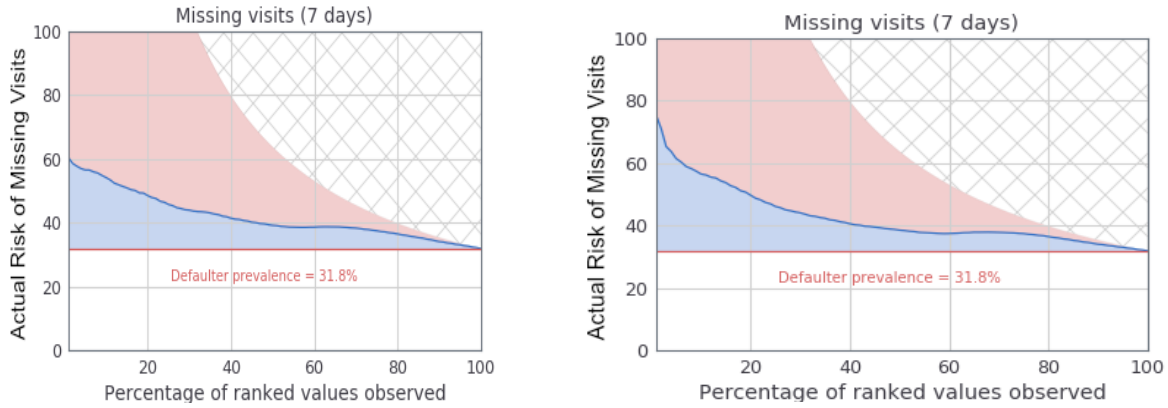
Results when Limiting Model Features

Next, we focused only on missed visits of 7 days or more, and turned our attention to optimizing the model. The first analysis explored the impact of reducing the number of variables used by the model. In Table 7 and Figure 5 we see the results of using fewer features. Figure 5a represents the model when restricted to 15 features, and Figure 5b represents the model restricted to 5 features (see Table 8 for list of variables). Surprisingly, we see an improvement when using fewer features. This warrants further investigation, though one likely explanation may be that the large number of features is causing the model to overfit on the training data.

Table 7: Missed appointment rate for all vs. "high risk" patients using different #s of variables

# days by which a visit was missed	Rate of missed appointments for all patients	Rate of missed appointments for the ML-predicted riskiest 10%	Rate of missed appointments for the ML-predicted riskiest 20%
7	31.8%	54.6% (15 features)	48.4 %
7	31.8%	57.4% (5 features)	49.61 %

Figure 5: Results when using top 15 features (5a) vs. top 5 features (5b)



5a 5b

Table 8: Feature set table

15 Features (Figure 5a)	5 Features (Figure 5)
Age Sex Occupation Number of household members Number of children Number of household members < 5 years Number of household members 5-14 years Number of partners at home Location Number of children tested Patient education level Number of pills per day Disclosed HIV status (yes/no) Availability of confidant (yes/no) Documented phone number	Age Number of household members Number of children Number of household members < 5 years Number of household members 5-14 years

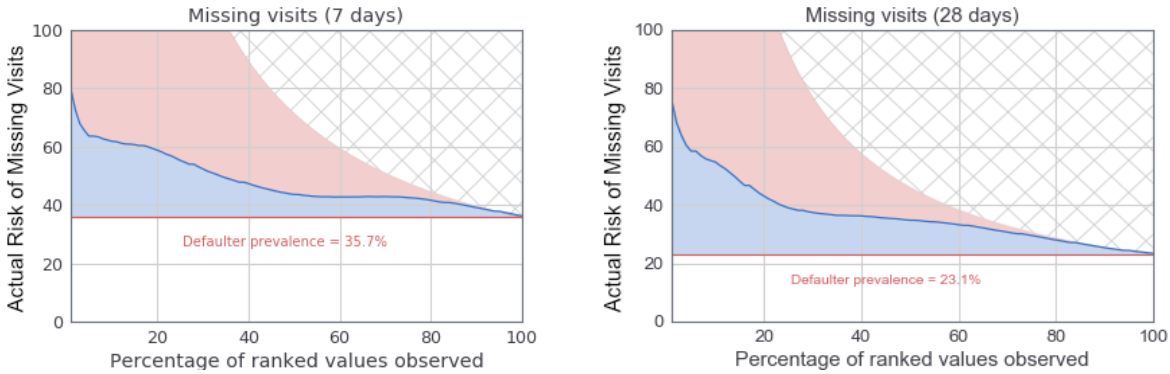
Results when Limiting Features and Sites

Finally, we built a model specifically for the Akumi site, which was the location with the largest number of patients and visits. Limiting the model to Akumi only and to the five features in Table 8, the model performed markedly better than the cross-site model (Table 9, Figure 6). Most notably, while looking at the top-20% of the risk scores, the model achieves 58.68% on PR-score. This suggests that location-specific models may be more effective than cross-site models, or that access to additional data (and cleaner data) may improve performance across sites.

Table 9: Akumi specific data

# of days by which a visit was missed	Rate of missed appointments for all patients	Rate of missed appointments for the ML-predicted riskiest 20%
7	35.7%	58.68%
28	23.1%	42.96%

Figure 6: Results when limiting to Akumi data and top 5 features



6a **6b**

Comparing the ML Algorithm with Simple Risk Stratification

As noted above, the ML algorithm was able to identify high-risk patients in each iteration of the model that was tested. In every case, the rate of missed appointments was higher for the subset of patients identified by the ML algorithm as "high risk" than it was for the group as a whole. However, it might be possible to achieve these results more simply, using a checklist or simple risk stratification approach.

To test this hypothesis, we assessed the default risk of patients who had not disclosed their HIV status and those who lacked a confidant; two of the characteristics that the activistas felt put patients at highest risk of default. As seen in Table 10, these variables by themselves and in combination did not predict risk of missing an appointment by 7 or more days; the risk of default was indistinguishable from that of the total patient population.

Table 10: Default risk for patients who have not disclosed their HIV status and/or do not have a confidant

Category	Risk of missing appt by ≥ 7 days
All patients (N=491)	31.8%
Non-disclosure (N=157)	30.6%
No confidant (N=225)	28.4%
Non-disclosure + no confidant (N=142)	31.7%

Section 4: Programming a Flag in the Infómovel User Interface

The Dimagi team developed a prototype workflow and illustrative user interfaces, described in detail in Appendix 3. Figure 7 illustrates the first screen in which an activista would see the flag.

Figure 7: Defaulter risk flag in Infómovel

Section 5: User Feedback

Despite the limited data available, the ML algorithm was able to identify subsets of patients at higher risk for default, and Dimagi successfully programmed a "risk flag" into the Infómovel user interface. However, in order for the initiative to be successful, health workers will have to use the flag to identify high-risk patients and to target specific interventions to this group. Thus, it is imperative to understand potential users' grasp of defaulting workflows, familiarity with patients, and familiarity with the existing application's defaulting workflows.

Towards this end, the Dimagi team obtained user feedback via a standard human-centered design approach. In September 2018, Dimagi visited two ICAP-supported health facilities - Muhala Expansao in Nampula district and Akumi Health Facility in Nacala district – to obtain insights from *activistas*, *gestors de casos*, and site-level adherence and psychosocial support (APSS) officers. The Dimagi team prioritized:

- Understanding Infómovel users' existing theories as to why patients missed visits
- Verifying that users' follow differentiated workflows for patients who have missed visits *versus* patients who do not
- Confirming that users are able to provide feedback on their current patient list regarding patients' adherence status, and therefore would have valuable inputs to provide as feedback to any predictor mechanism
- Assessing knowledge and acceptability of the current Infómovel application
- User testing of the icons for predicting who is at high risk of missing appointments
- Reviewing the application workflow with the "high risk" flag and feedback form

Details of the activities and responses can be found in Appendix 2, but in brief, key results included:

- There is a high level of user acceptance for Infómovel. *Activistas* agree that the application is useful to their main task of following up with patients
- Many *activistas* are not familiar with the defaulting workflows in the Infómovel application. This relates to the lack of use of the "Busca Activa" case list as well as the interpretation of the appointment-related icons in the main follow up case list.
- *Activistas* do not differentiate application workflows for active and defaulting patients, so do not grasp the utility of having a predictive flag
- In turn, this means that a risk calibration form will only be useful when implementers have clearly delineated the process for recording, tracking and a protocol defining the service package for 'at-risk' patients, using the Infómovel application

Conclusion

The ML algorithm was able to identify a subset of patients at higher risk of default, providing proof-of-concept for the project despite challenges related to the availability of data, and the need to manually import key variables from OpenMRS. Users were open to the idea of a 'flag' within Infómovel, and interested in learning more about how it would be used. A larger and cleaner dataset will improve the model, and alternative methods of aggregation (e.g., at the person level) as well as visit-specific features may prove to be beneficial for building future models and avoiding the challenge of overfitting due to a relatively small sample size. Development of high-risk workflows will also be critical to enable tailored interventions to be provided to patients at high risk of default.

Next Steps

The OpCon Mozambique pilot provided a foundation for future predictive work. Suggested next steps are described below; we note that activities 1-3 are already included in the Infómovel workplan for the current year.

1. OpenMRS and Infómovel Integration

Currently OpenMRS is the only source of medication pick-up and scheduled pick-up dates, and therefore the source used to classify patients as 'active', 'faltoso' or 'abandonos'. By integrating OpenMRS and Infómovel, we can have a better integrated dataset for further analysis and ensure data quality of key variables across digital collection systems. Luckily, Infomovel-OpenMRS integration functionality is expected to be included as an update of Infomovel app and will be available for CDC implementing partners in Mozambique during FY19; this will facilitate data collection across platforms to power an improved ML model.

2. Uploading Data on Missed Appointments to Infómovel

The critically important defaulter workflow is not being fully utilized in the Infómovel application. Uploading data on defaulting to Infómovel will help to keep the Busca Activa case list relevant and ensure that the activista users are sensitive to the differentiated workflows in the application. Working with activists on different patient statuses leading to different workflows will make the environment more receptive for future predictive risk flags.

3. Supervisory App

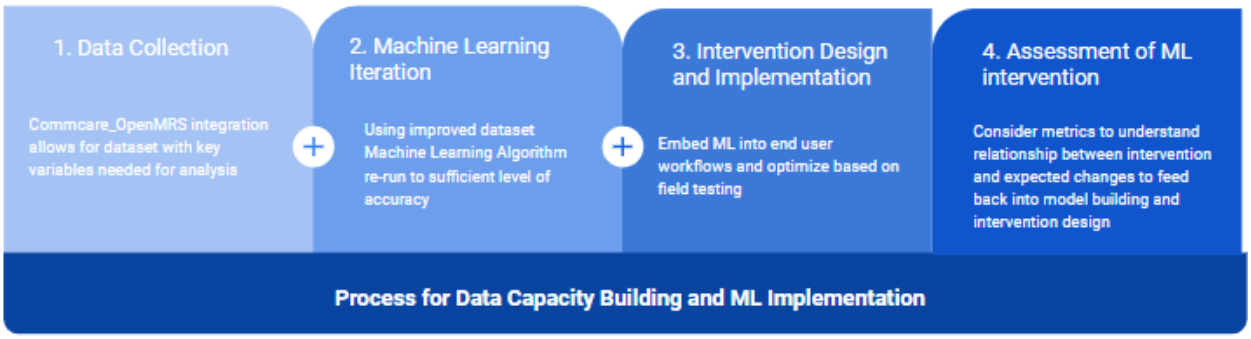
To better manage the app workflows and data management including tracking individuals who missed appointments, Dimagi is currently developing a supervisory app. This will help address consistency in following patients through the expected pathways in the application, and the data workflows on the back end of the application that would feed in to analyses.

4. Monitor Data Collection Stability to Re-Run ML Predictive Models

Once data collection and integration improve, new models can be tested to determine improvements in performance and to iterate on initial results. Following this, the aim would be to then design and field test a prototype UI for integration within the frontline health workflows, collaborating across organizational levels to determine requirements and test prototypes to determine optimal functionality.

A possible set of next steps to continue this work is illustrated in Figure 8:

Figure 8: Next steps



Appendix 1: Scoping Visit Activities

Date, location	Participants	Activity
21st May 2018 ICAP Maputo office	Maria Rein (CDC), Marcelo De Freitas (ICAP), Fernando Sitole, Nazera Nangy	• Introduction to the week's agenda • Sharing initial data reports • Briefing from the ICAP team re: how reports and data monitoring are used to monitor Infómovel usage in the facilities
22nd May 2018, Nampula DPS office	DPS (Provincial Ministry of Health)	• Introduction to Dimagi • Briefing of scoping agenda and purpose
22nd May 2018, Nampula. Mhuala Expansao Health Facility	Health Facility director of Mhuala Expansao	• Introduction to Dimagi • Briefing of scoping agenda and purpose

22nd May 2018, Nampula.
Mhuala Expansao Health Facility
	Case Manager (Celso Dava), APSS officer (Dr. Anibal), Health Facility focal point, Activistas	• Overview of the following workflows
		  ○ New patient enrolment on paper
		  ○ Upload to Infómovel and OpenMRS
		  ○ Allocation of new patients to activistas
		  ○ Process of first home visit
		  ○ Identifying patients who miss appointments
		  ○ Tracking patients who miss appointments
		• Discussion of:
		  ○ Number of patients who miss appointments
		  ○ Common reasons for defaulting
		  ○ Usability of the Infómovel application
		• Smaller group meeting with activistas
		  ○ Observing application navigation
		  ○ Observe usage of priority case lists
		  ○ Observe usage of defaulter case list

23rd May 2018	Activistas	• Accompanied activistas conducting home visits to observe how Infómovel was being used in real time

23rd May 2018	Case manager (Celso Dava), data officer (Narciso Ngalambe), APSS officer (Dr.Anibal)	• Detailed overview of Infómovel and OpenMRS use in
		  ○ Creating and uploading new patients from paper forms
		  ○ Updating treatment information
		  ○ Referral workflows
		  ○ Tracking patients who miss appointments

24th May 2018 Akumi Health Facility Nacala	Health Facility director	● Introduction to Dimagi ● Briefing of scoping agenda and purpose
24th May 2018 Akumi Health Facility Nacala	Case Manager (Diolfo Antonio) data office (Jaime Mangujo) and Dr Anibal	● Receiving an overview of the following workflows ○ New patient enrolment on paper ○ Upload to Infómovel and OpenMRS ○ Allocation of new patients to activistas ○ Process of first home visit ○ Identifying patients who miss appointments ○ Tracking patients who miss appointments ● Discussion on the following ○ Number of patients who miss appointments ○ Common reasons for defaulting ○ Usability of the Infómovel application
24th May 2018 Akumi Health Facility Nacala	Discussion with activistas	● Smaller group meeting with activistas ○ Observing application navigation ○ Observe usage of priority case lists ○ Observe usage of defaulter case list
24th May 2018 Akumi Health Facility Nacala	Case Manager (Diolfo Antonio) data office (Jaime Mangujo) and Dr Anibal	● Detailed overview of Infómovel and OpenMRS use in ○ Creating and uploading new patients from paper forms ○ Updating treatment information ○ Referral workflows ○ Tracking patients who miss appointments
28th May 2018, ICAP office, Maputo	Marcelo De Freitas, Fernando Sitole, Nazera Nangy	● Brief discussion on field findings and workflows ● Discussion options of OpenMRS access

Appendix 2: User Feedback Visit Activities

User Feedback Methods

The success of interventions targeting patients at risk of missing appointments depends on activistas' understanding of the reasons for default and the availability of a standardized protocol for at-risk patients. The user feedback process convened small groups of health workers to discuss the high-risk flag and its possible applications. No names were recorded, no audiotapes were made, and only aggregate responses are described.

In each facility, activistas were divided into groups of three. Each activista was asked to come up with as many reasons for default or missing appointments as s/he could think of. Once this was done, each group was tasked with classifying individual reasons into themes, and to present these to the group.

We then asked each group to describe the steps they current use to re-engage patients who have missed appointments back into care, using multiple scenarios (e.g., patient did not return because of medication side effects, patient forgot their appointment date, etc.).

To assess how thoroughly users were engaged with their patient list on Infómovel we had each activista note the following:

- The names of as many active patients as they could recollect (a random sample of these were checked against Infomovel case lists)
- Their recollection on whether they were currently defaulters or not
- Their opinion on which of their patients were at risk to default

In order to rapidly ascertain the level of clarity around Infómovel defaulter workflows, we asked *activistas* to rate their level of agreement with the following general statements:

- The application is useful to my work in following up with patients
- The application helps me understand which of my patients have defaulted in the past month
- The application helps me see which of my patients are at risk of defaulting
- It is important for me to know which of my patients are at risk of defaulting
- I think the defaulting icons in the app are easy to understand

Next, we shared the icon that is used in the test version of the application to flag patients at risk of default (see Appendix 3). Without letting the activistas know what it stood for, we asked them what they thought such an icon could indicate. Ideally, activistas would associate the icon with danger, or alert signs. We also asked users to quickly sketch an icon that they believed would be well suited to alerting users to patients at risk.

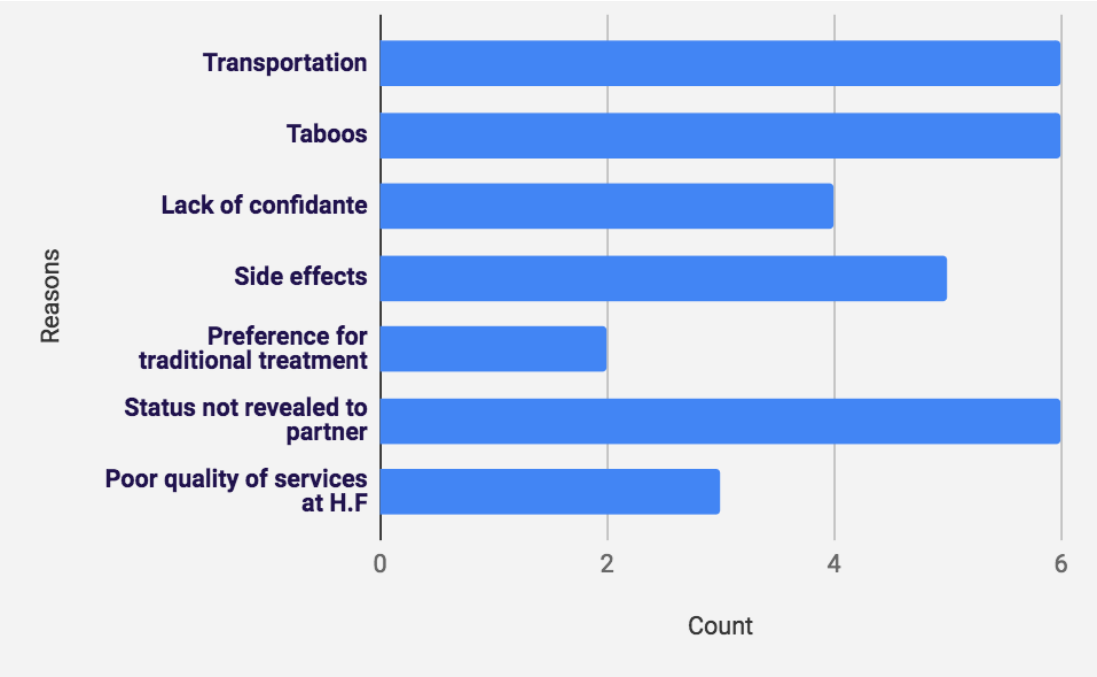
The final activity was a demonstration of the application with the defaulter risk flag and the form to calibrate the risk score (see Appendix 3).

User Feedback Results

Reasons for Default

The *activistas* reported the following reasons for patient default. The count in Figure A refers to the number of groups that cited each reason.

Figure A: Activistas' explanations for why patients default



Steps to Support Retention

Each group was able to identify multiple patient default scenarios to describe specific steps to bring patients back into care. Examples include:

Table A

Situation	Step 1	Step 2	Step 3	Step 4	Step 5
A patient has missed their visit date and refuses to continue treatment citing side effects	Explain the advantages of retention in care and raise awareness of side effects of medication	Provide examples on the effects of taking and not taking ART	Convince the patient to return to the health facility	Assign patient to Gestor (case manager)	Confirm with gestor the next day to see if the patient came for the consultation

A patient cannot be reached for a visit and has missed their visit date	Place phone call to patient to ascertain whereabouts	Alert the health facility so that the patient can be served	Talk to the patients about the importance about attending to the treatment and not missing consultations	Counsel patient about the advantage of returning to treatment	Wait for the patient to return and be allocated a health facility

Case List Familiarity

We found that activistas were able to recollect information regarding their active patients fairly quickly, which is encouraging for a future workflow that relies on this level of engagement. We asked for a minimum of 8 patients, in under 10 minutes, which almost all activistas were able to meet.

Knowledge of Application Workflows

The activistas expressed a largely positive attitude towards the existing Infómovel application, but the discussion revealed limited understanding of how defaulter tracking performed in the application.

For instance, there was unanimous agreement that Infómovel was useful to their job function. There was also initial unanimous agreement that the application was useful in tracking patients who miss appointments. When we probed into how the activistas were using the application for tracking, we realized that what they were really looking at was the icons for upcoming appointments in the follow-up module.

Similarly, even though the current version of the application has no flag for patients at risk of missing appointments, initially the majority of the activistas agreed that the application was useful for tracking risk. It took a few rounds of discussions before they were ready to state that the application was not useful in tracking patients at risk of default.

Icons

The meaning of the red exclamation point icon was not immediately clear to the activistas. This may not have a bearing on the final deployment, but indicates that more user testing would be required before deciding on the final format of the "flag" for patients at high risk of default in the user interface.

Application Demonstration

The final activity was a demonstration of the application with the defaulter risk flag and the form to calibrate the risk score. By doing this activity at the end we had time to really underscore all the concepts it touches upon - defaulting reasons, steps to retention, engagement with patients, icon recognition.

There were differences in both facilities in the grasp of the risk calibration form. We found that we had to dedicate most time in explaining to the activistas that although the risk flag would be developed at the back end, they still had a say in determining which of their patients were actually at-risk.

Appendix 3: Application Prototype

The following screenshots and diagrams represent the workflow for a patient after they have been identified as being at highest risk of missing an appointment.

Import patients via excel and confirm defaulter status

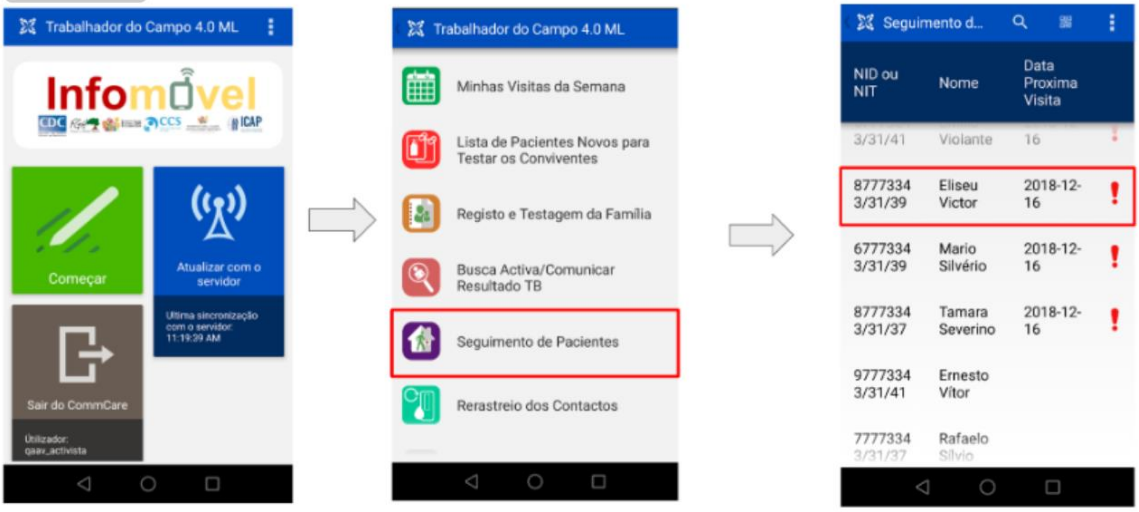
Step 1

defaulter_risk_score	nid	owner_id	nome	apelido	genero	data_ultima_consulta	data_proxima_consulta	... [more columns]
0	xx/xx/xxxxxx	1	ernesto	faustino	M	41333	41326	
1	xx/xx/xxxxxx	2	violeta	matine	F	41291	41305	
0	xx/xx/xxxxxx	3	benny	saide	M	41311	41319	
0	xx/xx/xxxxxx	4	larson	naimo	M	41320	41289	
0	xx/xx/xxxxxx	5	marven	passe	M	41317	41300	
1	xx/xx/xxxxxx	6	tamara	raul	F	41300	41311	
1	xx/xx/xxxxxx	7	miguel	pinho	M	41325	41321	
	xx/xx/xxxxxx	8	maria	bata	F	41304	41302	
0	xx/xx/xxxxxx	9	audrey	janota	F	41301	41293	

The **defaulter_risk_score** column has 1 = highest risk bracket to default. These cases will show up with an exclamation point next to them on the Activista case list.

The Activista then taps on their Follow-up Visits Module and views their case list

Step 2

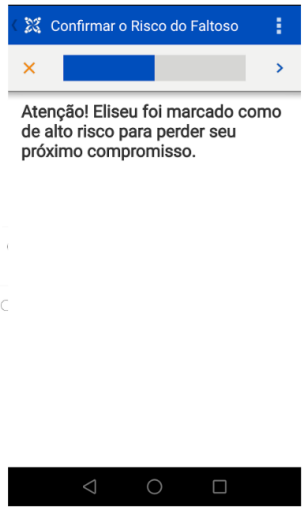


Step 3

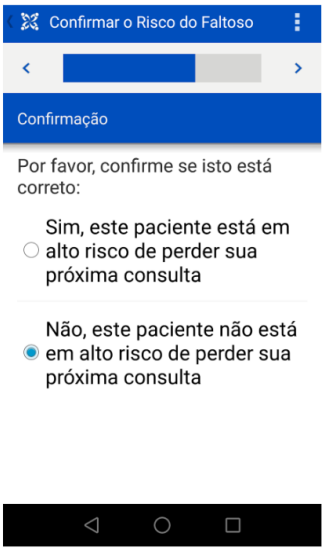
Only if the patient is high risk, the Activista will be directed to the Calibration Form.

- This form will allow the Activista to either confirm the patient is at high risk flag, or remove the flag.

Step 4

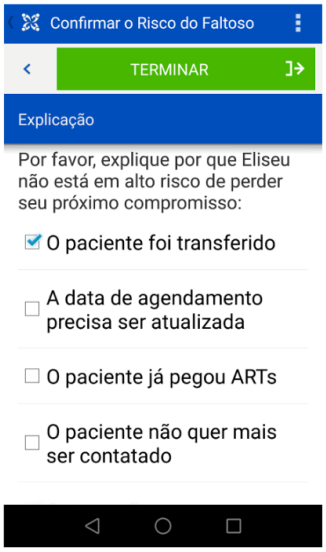


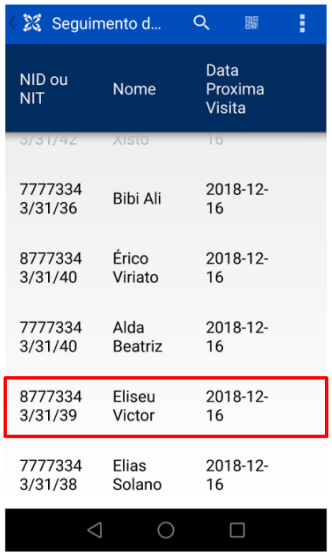
The Activista is given a warning about the risk to miss the next appointment



Step 5

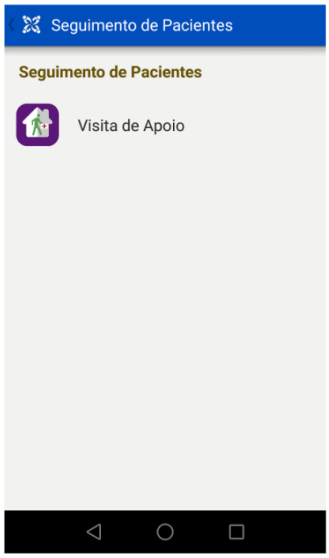
ML Workflow
- If NO is selected, the Activista will select a reason or add a new one not on this list

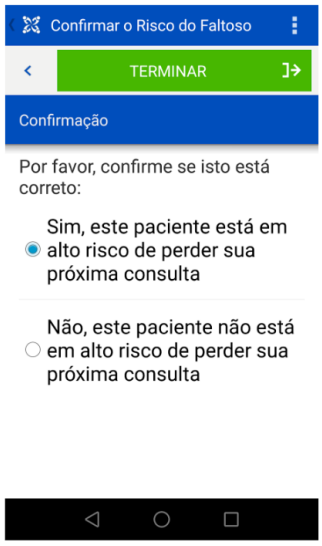




Step 6

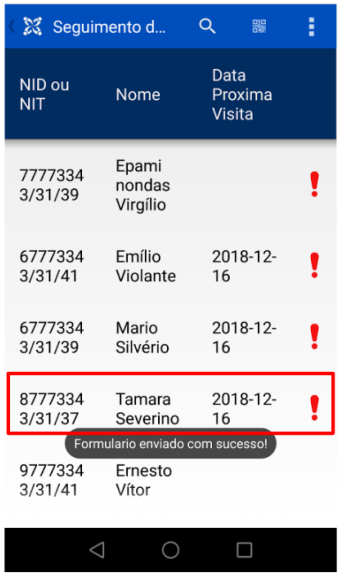
ML Workflow
- Returning to the case list, we now see that Eliseu is no longer flagged as high risk. The Activista is now guided to the Follow Up Visit Form





Step 7

In the case of a patient Tamara, If "yes" is selected by the Activista the exclamation point is retained in the case list view, and once the calibration is completed, the Activista is prompted to fill out the follow up form next



Step 8