ISGlobal - Barcelona Institute for Global Health

Delivering an Innovative Multi-disease Screening Tool to High-risk Migrant Populations (ISMiHealth)

STATISTICAL ANALYSIS PLAN

Version 2.0 September 16, 2024

Author/s:

Angeline Marie Cruz Vázquez, Pre-doctoral fellow, ISGlobal

Alba Cuxart Graell, Research Technician, ISGlobal

Aina Casellas, Biostatician, ISGlobal

Elisa Sicuri, Associate Research Professor, ISGlobal

Ana Requena-Méndez, Research Assistant Professor, ISGlobal

Tel: +34 93 227 54 00

Corresponding author: Ana Requena-Méndez, Research Assistant Professor, ISGlobal

Statistical Plan of the ISMiHealth study

Table of Contents

Abbreviations

TB	Tuberculosis
HIV	Human Immunodeficiency Virus
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
FGM	Female Genital Mutilation
PC	Primary Care
CDSS	Clinical Decision Support System
EPR	Electronic Patient Record
PCCs	Primary Care Centres
ICD-9	International Classification of Diseases Ninth Revision
ICD-10	International Classification of Diseases Tenth Revision
TST	Tuberculin Skin Test
IGRA	Interferon Gamma Release Assay
MBDS	Minimum Basic Dataset
IDIAP Jordi Gol	Fundació d'Investigació en Atenció Primària Jordi Gol i GurinaI
SAS	Andalusian Health Service
IT	Information Technology
SIDIAP	Sistema d'Informació per al Desenvolupament de la Investigació en Atenció-Primària
SISAP	Sistema d'Informació dels Serveis d'Atenció Primària
SD	Standard Deviation
IQR	Interquartile Range
OR	Odds Ratio
aOR	Adjusted Odds Ratio
PCT	Primary Care Team
MEDEA Index	Index of Socioeconomic Deprivation
EAP	Equipo de Atención Primaria
CAP	Centro de Atención Primaria
AST	Aspartate Aminotransferase
ALT	Alanine Aminotransferase
GGT	Gamma-Glutamyl Transpeptidase
STDs	Sexually Transmitted Diseases
ACT	Anatomical Therapeutic Chemical

1. Introduction

1.1 Preface

Migration is a significant, complex, and growing global phenomenon of critical importance to European countries[1]. Migrants may have particular health needs; they are disproportionately affected by key infections, including tuberculosis (TB), human immunodeficiency virus (HIV), hepatitis B virus (HBV) and hepatitis C virus (HCV) in Europe; and this has been seen associated with the incidence in the country of origin, socio-economic risk factors and other comorbidities on arrival to the host Europe[2]. Similarly, certain imported diseases, only prevalent in migrant populations[3-6], are at risk of transmission in non-endemic areas under special circumstances (through transplants, blood-transfusion or congenitally)[5] and are potentially severe, particularly if there is a diagnosis delay, in immunosuppressed patients[6]. Finally, female genital mutilation (FGM) is a neglected condition, with estimates about 38000 migrant women in some European countries[7].

The new guidelines from the European Centre for Disease Control and Prevention[8] on infectious diseases screening and vaccination in new-arriving migrant populations and alongside country-specific guidelines, promote screening implementation at the primary care (PC) level[9]. In particular, ECDC guidelines call for innovative approaches to multi-disease testing and vaccination in high-risk migrants[8]. With the aim of improving patient care by strengthening medical decisions, there has been an outstanding development of clinical decision support systems (CDSS) in the last decade[10]. In such tools, the characteristics of an individual patient are matched to a computerized clinical staff knowledge base in patient-specific assessments and recommendations are then presented in the electronic patient record (EPR) system to the clinical staff for decision. Our research group has already developed an innovative prototype digital software (CRIBMI)[11]. Data shows the digital tool is feasible to implement, acceptable to healthcare professionals and migrants, and well adapted to the PC setting[12], with data suggesting that the digital tool is improving the diagnosis of imported conditions but also other common communicable diseases in migrants when comparing it with a standard training programme[12].

1.2 Purpose of the analysis

The purpose of this analysis is to evaluate the health impact of the ISMiHealth tool validating the tool in a higher number of primary care centres (PCCs) in Catalonia, Spain; and to explore preliminary effectiveness data of the tool in another healthcare setting, Andalusia, Spain.

2. Study objectives and Endpoints

2.1 Main Study Objectives

'To develop an integrated and PC based, communicable diseases and FGM screening programme for migrant populations in Spain in order to promote better health and integration.'

For this, the research group will evaluate the health impact of the ISMiHealth tool in the site of Catalonia, while preliminary effectiveness data will be evaluated in the site of Andalusia.

2.2 Specific Study Objectives

Primary objective 1: To compare the detection rate per month of all aggregated infections (HIV, HBV, HCV, TB, *T.cruzi*, *S.stercoralis* and *Schistosoma spp.* infections) between the intervention and control centres over at least five years until the end of the intervention. In the Andalusian site, syphilis, latent TB and intestinal parasites will also be included.

Secondary objective 2: To compare the detection rate per month of each individual condition, the infections and FGM cases, between the intervention and control centres over at least five years until the end of the intervention.

Secondary objective 3: To compare the number of early HIV diagnoses between the intervention and control centres over at least five years until the end of the intervention.

Secondary objective 4: To compare the number of early HBV and HCV diagnoses between the intervention and control centres over at least five years until the end of the intervention.

Secondary objective 5: To compare the number of screening tests performed for all aggregated infections between the intervention and control centres over at least five years until the end of the intervention.

Secondary objective 6: To compare the number of screening tests performed for each individual condition between the intervention and control centres over at least five years until the end of the intervention.

Secondary objective 7: To estimate associations between the screening performance of all aggregated infections with different socio-demographic and epidemiological factors.

Secondary objective 8: To estimate associations between the screening performance of each individual condition with different socio-demographic and epidemiological factors.

Secondary objective 9: To measure the number of the diagnosed individuals with follow-up visits in the hospital of reference.

Secondary objective 10: To measure the number of the diagnosed individuals that have received treatment.

Secondary objective 11: To estimate the prevalence of each infection and FGM in the migrant population.

2.3 Endpoints

Primary Endpoints

Endpoint 1: For the detection rate of all aggregated infections, the monthly detection rate based on positive serologies, chest radiographies, the International Classification of Diseases Ninth Revision (ICD-9 for Andalusia) or Tenth Revision (ICD-10 for Catalonia) of FGM and/or gynaecologist referrals will be considered within the migrant patients who visited their assigned centre during the intervention period. Also, positive tuberculin skin test (TST) and/or Interferon Gamma Release Assay (IGRA) and stool samples will be considered for the Andalusian site. Control and intervention PCCs will be compared before and after the implementation.

Secondary Endpoints

Endpoint 2: For the detection rate of each individual condition, the infections and FGM, the monthly detection rate based on positive serologies, chest radiographies, ICD-9/ICD-10 of FGM and/or gynaecologist referrals, TST/IGRA and stool samples (for Andalusia) will be considered within the migrant patients who visited their assigned centre during the intervention period. Control and intervention PCCs will be compared before and after the implementation.

Endpoint 3: The number of early diagnoses of HIV will be assessed using the CD4 cell count of migrant patients diagnosed with an HIV infection, and will be compared between the control

and intervention PCCs before and after the implementation. If possible, we will estimate the monthly detection rate of early HIV diagnoses and also compare it between the intervention and control PCCs.

Endpoint 4: The number of early diagnoses of HBV and HCV will be assessed using the levels of transaminases, platelets, bilirubin, and clotting parameters from blood analyses of migrant patients diagnosed with HBV or HCV, and will be compared between the control and intervention PCCs before and after the implementation. If possible, we will estimate the monthly detection rate of individuals diagnosed with HBV or HCV presenting high levels of transaminases, platelets, bilirubin, and clotting parameters and also compare it between the intervention and control PCCs.

Endpoint 5: The number of the screening tests performed of all aggregated infections (serological tests, chest radiographies, TST or IGRA tests and stool samples in the case of the Andalusian site) will be compared between the intervention and control PCCs before and after the implementation.

Endpoint 6: The number of the screening tests performed for each individual condition will be compared between the intervention and control PCCs before and after the implementation.

Endpoint 7: Associations between the screening performance of the aggregated infections with different socio-demographic and epidemiological factors, such as, sex, age, geographic areas immunosuppression status, being attended in an intervention centre, fulfilling the screening criteria, among others, will be analysed.

Endpoint 8: Associations between the screening performance of each individual condition with different socio-demographic and epidemiological factors, such as, sex, age, geographic areas immunosuppression status, being attended in an intervention centre, fulfilling the screening criteria, among others, will be analysed.

Exploratory Endpoints

Endpoint 9: The number of migrant patients under follow-up in the hospital after a diagnosis of an infection or FGM will be estimated using the Minimum basic dataset (MBDS).

Endpoint 10: The number of migrant patients under treatment after a diagnosis will be estimated using the data from the EPR systems and MBDS.

Endpoint 11: The prevalence of each condition will be estimated using as the denominator the total number of tested migrant patients for that specific condition.

3. Study Methods

3.1 Study design

A pragmatic cluster controlled randomised trial will be conducted in 35 PCCs of Catalonia to explore and assess the effectiveness of the digital tool ISMiHealth. Additionally, a pilot clustered controlled randomised trial will be conducted in Andalusia, implementing the ISMiHealth tool in six PCCs to explore preliminary effectiveness data. This user-friendly innovative digital tool (ISMiHealth) that aims to facilitate targeted screening to recently arrived migrants presenting for a routine appointment, will be integrated into the local EPR system of the intervention-PCCs of the two different settings of Spain (Catalonia and Andalusia). The tool will send real-time recommendations to health professionals concerning the screening of infectious diseases and female genital mutilation based in the individual risk of each migrant patients, using the

variables sex, age, and country of origin. Meanwhile, the control-PCCs will follow the routine care practice. All centres will receive training sessions on Migrant Health and Female Genital Mutilation.

The screening programme will be implemented in a real-life context in the selected PCCs, and the receivers/users of the tool will be all health professionals (general practitioners and nurses) working at those centres. On the other hand, the tool and the outcomes of the study will be evaluated in the migrant patients, the indirect beneficiaries of the tool.

3.2 Inclusion and exclusion criteria

Inclusion criteria of PCCs:

- Centres with a migration density higher than 7%

Inclusion criteria of PC professionals:

- Aged ≥18 years old working at the selected PCCs

Inclusion criteria of migrant populations:

- Individuals assigned to a PCC
- Patients attending a PCC for any reason
- Aged ≥15 years old in the Catalonian site
- Aged ≥14 years old in the Andalusian site
- Coming from Africa, Latin-America, Asia and Eastern Europe following the categorization of the UN Statistical Commission[13]

Exclusion criteria:

- For the active TB recommendation, migrants residing in the host country for more than five years.
- For FGM recommendation, being a male.

3.3 Health centres selection, randomisation and masking

Regarding the selection of the participating PCCs in Catalonia, the Fundació d'Investigació en Atenció Primària Jordi Gol i Gurina (DIAP Jordi Gol) research team will find areas where there are referents with an interest in participating in the study. The list of PCCs belonging to these areas, where the referral to the reference laboratory is more feasible, will be provided. With this final list, we will proceed with the randomization stratified by area and by density of migrant population in the area (low, medium and high). For the selection of the PCCs in Andalusia, an IT committee of the Andalusian Health Service (SAS) will provide a list of the possible centres that could participate in the study and the randomization will be carried out using the same stratification method mentioned previously.

The randomization will be performed through a statistical software for each pair of selected health centres stratified by study area. Therefore, for each pair of PCCs, one PCC will be randomly selected to implement the screening programme through the ISMiHealth software, and it will be compared with the other PCC of the same area where the PCCs do follow the current practices in the routine care. In both cases, health professionals will receive a training session on migrant health; the training contents will include for each condition, epidemiological aspects, diagnosis, treatment and the screening recommendation.

3.4 Study Variables

Data will be collected retrospectively, therefore, the data for all the variables mentioned in Annex 1 will be collected after the year of intervention is completed.

4. Sample size calculation

Considering that all PCCs included in the study have a migration density higher than 7%, we have estimated that by including 32 or more centres in Catalonia we will achieve a difference in difference detection rate of all aggregated infections between intervention and control centres of 2, with a 95% precision. A comparative analysis of the PCCs, in relation to the migration density, will be first performed to select the pairs of centres; then, each pair of selected health centres will be randomized, stratifying by study area, using the Statistical software Stata. To assure the validity of results the intervention centres (where the ISMiHealth tool will support the clinicians' decisions) will be compared to the control centres (where the routine care will be followed) with more similar characteristics.

41 PCCs will participate, 35 in the Catalonian site and six in the Andalusian site. In Catalonia 18 PCCs will be intervention and 17 control, not pairing one of the centres; while in Andalusia there will be 3 intervention and 3 control PCCs. The users of the ISMiHealth tool will be the health professionals of the intervention centres and as participants of the study, they will decide whether to follow the recommendations to screen migrant patients, the indirect beneficiaries of the tool. Approximately 980 health professionals will participate in the study.

5. General Considerations

5.1 Timing of Analyses

In the Catalan Institute of Health, data will be extracted retrospectively after the year of implementation from the EPR system (eCAP) by the Information Technology (IT) staff of the Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP). The study variables will be pseudo-anonymized, with technical and functional separation, by a SIDIAP data manager. SIDIAP will be responsible for the data processing in Catalonia. Regarding SIDIAP source data, the strict SIDIAP security standards will be followed.

In Andalusia, the data will be extracted retrospectively after the year of implementation by a specialized IT committee of the Andalusian Health Service Institution of the area involved in this project (Almería). The information will be extracted retrospectively and will include: 1) the check-list from the tool and 2) the demographic and clinical variables. Then, the IT committee will link this information with the laboratory data of the reference hospital (Hospital de Poniente in Almeria) and will pseudo-anonymize the data.

Thereafter, the biostatistician responsible for the data analysis will have access to the pseudo-anonymized data. The pseudo-anonymized data, which will be analysed after the year of intervention is completed, will include the variables detailed in Annex 1 for the targeted migrant population of the PCCs of Catalonia and Andalusia included in the study.

In any case, the data from the Catalonian centres will be analysed separately from data of the PCCs from Almeria where a process evaluation will be performed and only preliminary effectiveness data will be estimated.

5.2 Covariates

It is presumed that the heterogeneous composition of the migrant population by country of birth within the study sites, and the type of centre (urban/rural), might have an influence on the impact and performance of the ISMiHealth tool. It is hypothesised that a higher detection rate and screening performance for specific conditions will be associated with migrants coming from endemic areas and/or high prevalent areas, as well as regarding the type of centre.

5.3 Missing data

A limitation of the retrospective data extraction is the possibility of a sizable number of missing values. All available data of the variables will be described and the proportions of missing values will be reported. If the percentage of missing data is high ($\geq$50%), listwise deletion -deleting observations with more than one missing value- or dropping variables -discarding variables with a big amount of missing data- will be carried out.

5.4 Interim Analysis and Data Monitoring

No interim analysis will be conducted. However, the Sistema d'Informació dels Serveis d'Atenció Primària (SISAP) will supervise the technical monitoring of the tool in the implementation phase of the study. Simultaneously, aggregated data will be obtained via SIDIAP during the process of implementation of the tool in the Catalan Institute of Health with the sole objective of technical monitoring to identify alert errors and provide subsequent remediation.

6. Data analysis

Since the ISMiHealth tool is in different phases of implementation in each site, and the algorithms are based on different recommendations, there will be two independent data analyses: i) in Catalonia, the health impact of the tool will be evaluated in 35 PCCs; and ii) in Andalusia, the preliminary effectiveness data will be estimated in the six PCCs where the tool will be pilot tested.

The precision for the reporting of the results will be 95%, considering statistically significant p.values<0.05.

First, for the description of the sample and data obtained from the screening program in each study site, summary statistics will be presented as frequencies for categorical variables, as means with standard deviations (SD) for normally distributed continuous variables and medians with interquartile range (IQR) for non-normally distributed continuous variables (Table 1 and 2).

Data analysis in Catalonia

a. **Health impact of the tool**. The effectiveness of the tool will be evaluated and compared between the intervention and control centres. Significant deviations between intervention and control centres before the implementation of the screening program will be suggestive of the main underlying assumption, presenting parallel trends prior to implementation.

- **The primary outcome** of the study will be the detection rate of the aggregated infections, while the detection of the individual conditions will be considered as **secondary outcomes of the study**. To analyse the effect of the intervention on the outcomes, difference in differences approach will be performed using a generalized linear model. Intervention and control PCCs will be compared before and after implementation with respect to the monthly detection rate. Standard errors will be clustered at the

intervention level. Linear regressions will be performed for the monthly detection rates of the aggregated infections, and of each individual condition, comparing the intervention and control centres before and during the intervention period.

- **Other secondary outcomes** will be the proportion of screening tests performed for the aggregated infections and for each individual condition, and factors associated with having a higher screening performance. For this, associations with continuous variables will be evaluated using t-tests, or one-way analysis of variance (one-way ANOVA) for normally distributed quantitative variables, while for not normally distributed quantitative variables, Wilcoxon Rank-Sum and Kruskal-Wallis tests will be carried out. Pearson's chi-square test or Fisher's exact test will be performed to evaluate the associations between categorical variables. In addition, odds ratio (OR) will be estimated (Table 3). Mixed effects logistic regression models will be carried out to identify associations between the screening performance and socio-demographic and epidemiological characteristics, and adjusted odds ratio (aOR) will be estimated, using the region of the PCC as a random intercept (Table 3).

b. The **prevalence** of each individual condition will be estimated and reported with its respective 95%CI. The prevalence for each infection was defined as the number of diagnoses identified by the screening tests performed among the people that were tested. Serology tests will be contemplated for HIV, HBV, HCV, *T.cruzi*, *S.stercoralis* and *Schistosoma spp.* infections; for active TB, chest radiographies will be examined; for FGM, the completed questionnaires and/or referrals to specialists will be considered.

See table 4 for details of the methodology employed for the outcomes.

Data analysis in Andalusia

A similar analysis approach will be carried out in Andalusia but the limited number of PCCs may prevent to estimate the health impact of the tool, therefore, we plan to estimate preliminary effectiveness data. In addition, the analysis for syphilis, latent TB and intestinal parasites will be executed. Nevertheless, because the pilot study will be conducted in six PCCs of one area (Almería), we won't be able to carry mixed-effects logistic regression models and instead, multiple logistic regressions will be performed.

7. Tables and Figures

Table 1. Description of the ISMiHealth sample.

Variables	Summary statistics
PC professional	
Primary care team (PCT)	N (%)
Type of health professional	N (%)
Professional category	N (%)
Sex	N (%)
Age	Mean (SD)
Centre	
Region	N (%)
Type of centre (urban or rural)	N (%)
Index of socioeconomic deprivation (MEDEA index)	Mean (SD) or Median (IQR)
Number of professionals (general practitioners and nurses)	Mean (SD) or Median (IQR)
Migration density	Mean (SD) or Median (IQR)

Number of migrants assigned to the centres	Mean (SD) or Median (IQR)
Number of migrants who did not visit the centres during the intervention	Mean (SD) or Median (IQR)
Number of migrants with at least one visit during the intervention period	Mean (SD) or Median (IQR)
Arm of the intervention	N (%)
Patient	
Sex	N (%)
Age	Mean (SDI)
Area of birth	N (%)
Assigned PCT	N (%)
Assigned health professional	N (%)
Time registered in the healthcare system	Mean (SD) or Median (IQR)
Number of visits	Mean (SD) or Median (IQR)
Immunosuppression status	N (%)
Migrants with screening criteria for HIV	N (%)
Migrants with screening criteria for syphilis^	N (%)
Migrants with screening criteria for HBV	N (%)
Migrants with screening criteria for HCV	N (%)
Migrants with screening criteria for active TB	N (%)
Migrants with screening criteria for latent TB^	N (%)
Migrants with screening criteria for a *T.cruzi* infection	N (%)
Migrants with screening criteria for a *S.stercoralis* infection	N (%)
Migrants with screening criteria for a *Schistosoma spp.* infection	N (%)
Migrants with screening criteria for intestinal parasites^	N (%)
Migrants with screening criteria for FGM	N (%)

PC: primary care, PCT: primary care team, HIV: human immunodeficiency virus, HBV: hepatitis B virus, HCV: hepatitis C virus TB: tuberculosis, FGM: female genital mutilation, SD: standard deviation, IQR: interquartile range; N = frequency

Table 2. Data obtained throughout the ISMiHealth screening program.

Variables	Summary statistics
Migrants tested for HIV	N (%)
Migrants tested for syphilis^	N (%)
Migrants tested for HBV	N (%)
Migrants tested for HCV	N (%)
Migrants tested for active TB	N (%)
Migrants tested for latent TB^	N (%)
Migrants tested for a *T.cruzi* infection	N (%)
Migrants tested for a *S.stercoralis* infection	N (%)
Migrants tested for a *Schistosoma spp.* infection	N (%)
Migrants tested for intestinal parasites^	N (%)
Migrants screened for FGM	N (%)
Migrants diagnosed with HIV	N (%)
CD4 cell count	Mean (SD) or Median (IQR)
HIV viral load	Mean (SD) or Median (IQR)
Migrants diagnosed with syphilis^	N (%)

Migrants diagnosed with HBV	N (%)
Migrants diagnosed with HCV	N (%)
Migrants diagnosed with active TB	N (%)
Migrants diagnosed with latent TB^	N (%)
Migrants diagnosed with *T.cruzi* infection	N (%)
Migrants diagnosed with *S.stercoralis* infection	N (%)
Migrants diagnosed with *Schistosoma spp.* infection	N (%)
Migrants diagnosed with intestinal parasites^	N (%)
Laboratory parameters	Mean (SD) or Median (IQR)
Migrants diagnosed with FGM	N (%)
Referrals to specialists	N (%)
Other diagnoses	N (%)
Comorbidities	N (%)
Drugs prescribed during the intervention	N (%)
Medication adherence (in days)	Mean (SD) or Median (IQR)
Hospitalised patients	N (%)
Diagnoses during hospitalisation	N (%)
Time hospitalised	Mean (SD) or Median (IQR)
Patients who needed Intensive Care Unit	N (%)
Drugs prescribed during hospitalisation	N (%)
Medication adherence of drugs prescribed during hospitalisation (in days)	Mean (SD) or Median (IQR)
Follow-up visits after hospitalisation	Mean (SD) or Median (IQR)
Diagnoses made during follow-up visits after hospitalisation	N (%)
Drugs prescribed during follow-up visits after hospitalisation	N (%)
Medication adherence of drugs prescribed in follow-up visits after hospitalisation (in days)	Mean (SD) or Median (IQR)
Cured patients	N (%)
Patients who died during the intervention period	N (%)

HIV: human immunodeficiency virus, HBV: hepatitis B virus, HCV: hepatitis C virus, TB: tuberculosis, FGM: female genital mutilation, SD: standard deviation, IQR: interquartile range; N = frequency

Table 3. Associations of different factors with XX tests performed

Variables	XX test performed n/N (%)	OR (95% CI)	p. value	aOR (95% CI)	p. value
Centre					
Type of centre (urban or rural)					
Index of socioeconomic deprivation (MEDEA index)					
Arm of the intervention					
Patient					
Sex					
Age					
Area of birth^					
Assigned PCT					

Referrals to specialists					
Drugs prescribed during the intervention					
Other diagnoses					
Immunosuppression status					
Migrants with screening criteria for XX^					
CD4 cell count, viral load and/or laboratory parameters*					
Hospitalisation					
Follow-up visits after hospitalisation					
Mortality					

PCT: primary care team, XX: aggregated infections or any of the conditions, OR: odds ratio, aOR: adjusted odds ratio, CI: confidence interval; Region will be used as a random intercept for the logistic regressions; ^Only one of these variables will be added to the multiple logistic regression model due to highly correlation; *When applicable

Table 4. Description of the methodology employed for each outcome

Outcome		**Methodology**
Detection rate of all aggregated infections		Difference in differences analysis
Detection rate of each individual condition	HIV cases	Difference in differences analysis
	Syphilis cases^	
	HBV cases	
	HCV cases	
	active TB cases	
	latent TB cases^	
	T.cruzi infection cases	
	S.stercoralis infection cases	
	Schistosoma spp. infection cases	
	Intestinal parasites cases^	
	FGM cases	
Early diagnoses of infections	HIV	Bivariate analysis and, if possible, difference in differences analysis
	HBV	
	HCV	
Screening tests performed of all aggregated infections		Bivariate analysis
Screening tests performed for each individual condition	HIV	Bivariate analysis
	Syphilis^	
	HBV	
	HCV	
	active TB (chest radiographies)	
	latent TB^ (TST/IGRA)	
	T.cruzi infections	
	S.stercoralis infections	
	Schistosoma spp. infections	

	Intestinal parasites^	
	FGM (completed questionnaires)	
Associations with the screening performance for all aggregated infections		Bivariate and multivariate analysis
Associations with the screening performance for each individual condition	HIV	Bivariate and multivariate analysis
	Syphilis^	
	HBV	
	HCV	
	active TB	
	latent TB^	
	T.cruzi infections	
	S.stercoralis infections	
	Schistosoma spp. infections	
	Intestinal parasites^	
	FGM	
Patients under follow-up after diagnosis		Bivariate analysis
Patients under treatment after diagnosis		Bivariate analysis
Cured patients		Bivariate analysis
Prevalence of each condition	HIV	Numerator: number of detected cases for each specific condition. Denominator: number of tested patients for each specific condition.
	Syphilis^	
	HBV	
	HCV	
	active TB	
	latent TB^	
	T.cruzi infections	
	S.stercoralis infections	
	Schistosoma spp. infections	
	Intestinal parasites^	
	FGM	

HIV: human immunodeficiency virus, HBV: hepatitis B virus, HCV: hepatitis C virus, TB: tuberculosis, FGM: female genital mutilation, TST: tuberculin skin test, IGRA: Interferon Gamma Release Assay; ^Only in the Andalusian site; The detection rate of all aggregated infections and the individual conditions will be calculated using the numbers of diagnoses per month over at least five years until the end of the intervention.

Figure 1. Flow chart of the migrant population

Figure 2. Monthly detection rates of the aggregated infections

Figure 3-9. Monthly detections of each individual condition (infections and FGM)

8. Reporting Conventions

Dates will be reported using the format MM/DD/YYYY. P-values will be reported to 3 decimal places. P-values less than 0.001 will be reported as '<0.001'. Statistics such as means, standard deviations, median values and percentages, will be calculated using the same decimal digits as the raw data. Rounding off will be performed on the final calculation value and reported according to publication requirements.

9. Technical Details

The data of the ISMiHealth study will be analysed with Stata Statistical Software: Release 16. At the time of the writing, the operating system of the computer that will be used for the analysis is Windows.

10. Summary of Changes to the Protocol

The protocol version 1.0 (March 10, 2022) has been approved by the institutional review boards at Hospital Clínic (06/16/2022, HCB/2022/0363), Clinical Research Ethics Committee of the Primary Care Research Institute IDIAPJGol (06/22/2022, 22/113-P) and the Almería Research Ethics Committee (07/27/2022, EMC/apg).

No substantive changes to the methods by which this study is to be conducted have been submitted.

11. References

1. Willekens F, et al. International migration under the microscope. Science (80-). 2016;352(6288):897 -9.

2. European Centre for Disease Prevention and Control (ECDC). Assessing the burden of key infectious diseases affecting migrant populations in the EU/EEA. Stockholm. 2014. doi:10.2900/28792. 2014.

3. Asundi A, et al. Prevalence of strongyloidiasis and schistosomiasis among migrants: a systematic review and meta-analysis. Lancet Glob Heal 2019;7(2):e236 -48.

4. Requena-Méndez A, et al. Prevalence of Chagas Disease in Latin-American Migrants Living in Europe: A Systematic Review and Meta-analysis. PLoS Negl Trop Dis. 2015;9(2):1 - 15.

5. Requena-Méndez A, et al. Health Policies to Control Chagas disease transmission in European Countries. PLoS Negl Trop Dis. 2014;8(10).

6. Requena-Méndez A, et al. Evidence-based guidelines for screening and management of strongyloidiasis in non-endemic countries. Am J Trop Med Hyg. 2017;97(3):645 -52.

7. Jordal M, et al. Challenges in providing quality care for women with female genital cutting in Sweden - A literature review. Sex Reprod Healthc [Internet]. 2018;17(September 2017):91 -6. Available from: https://doi.org/10.1016/j.srhc.2018.07.002

8. European Centre for Disease Prevention and Control (ECDC). Public health guidance on screening and vaccination for infectious diseases in newly arrived migrants within the EU/EEA. 2018.

9. Public Health England. Migrant health guide. Available from: www.gov.uk/government/organisations/public-health-england

10. Van Dort BA, et al. Optimizing clinical decision support alerts in electronic medical records: a systematic review of reported strategies adopted by hospitals. J Am Med Informatics Assoc 2020. 2021 Jan 15;28(1):177-183.

11. Requena-Méndez A. Migrant Health Conference -ISTM. In: CRIBMI: Improving screening strategies for migrants in Primary Care (PC). 2018.

12. Sequeira-Aymar E, Cruz A, Serra-Burriel M, di Lollo X, Gonçalves AQ, Camps-Vilà L, et al. Improving the detection of infectious diseases in at-risk migrants with an innovative integrated multi-infection screening digital decision support tool (IS-MiHealth) in primary care: A pilot cluster-randomized controlled trial. J Travel Med. 2021;1–11.

13. United Nations Statistical Division. Geographic Regions [Internet]. 2023 [cited 2023 Feb20]. Available from: https://unstats.un.org/unsd/methodology/m49

12. Annex

Annex 1. Variables of the ISMiHealth study

Variable name	Definition	Variable type	Values
PC professional (general practitioners and nurses)			
pct_name	Name of the team to which the professional belongs	Qualitative	PCTs in Catalonia: Equipo de Atención Primaria (EAP) Gavà-1 EAP Esparreguera EAP EAP Molí Nou EAP Gornal EAP Sant Andreu de la Barca EAP Gavà-2 EAP Camps Blancs EAP Sant Les Planes EAP Can Vidalet EAP El Castell EAP Florida Sud EAP Florida Nord EAP Lleida Rural Sud EAP Alfarràs - Almenar EAP Ponts EAP Almacelles EAP Artesa de Segre EAP Alcarràs EAP El Morell EAP Tàrraco EAP Sant Pere i Sant Pau EAP La Canonja/Bonavista EAP Constantí EAP Jaume I EAP Sant Salvador/Els Pallaresos EAP Torreforta EAP El Salou EAP Deltebre EAP Gandesa EAP Baix Ebre EAP Amposta EAP Sant Carles de la Ràpita EAP El Temple EAP L'Ametlla de Mar – El Perelló EAP Ulldecona – La Sènia PCTs in Andalusia: EAP Aguadulce Sur EAP Adra EAP El Ejido Norte

			EAP Roquetas Sur EAP la Mojonera EAP Puebla de Vícar
pc_hp	Type of health professional	Qualitative	General practitioner/Nurse
pc_category	Professional category	Qualitative	Professionals in training (residents), professionals on staff (permanent), professionals with intermediate and advanced vocational training
pc_sex	Sex of the PC professional	Qualitative	Female/Male
pc_age	Age of the PC professional	Quantitative	18-70 years old
Centre			
pcc	PCC where the patient is attended	Qualitative	PCCs in Catalonia: Centro de Atención Primaria (CAP) Gavà-1 CAP Esparreguera CAP Molí Nou CAP Gornal CAP Sant Andreu de la Barca CAP Gavà-2 CAP Camps Blancs CAP Sant Les Planes CAP Can Vidalet CAP El Castell CAP Florida Sud CAP Florida Nord CAP Lleida Rural Sud CAP Alfarràs CAP Almenar CAP Ponts CAP Almacelles CAP Artesa de Segre CAP Alcarràs CAP El Morell CAP Tàrraco CAP Sant Pere i Sant Pau CAP La Canonja CAP Bonavista CAP Constantí CAP Jaume I CAP Sant Salvador CAP Els Pallaresos CAP Torreforta CAP El Salou CAP Deltebre CAP Gandesa CAP Baix Ebre CAP Amposta CAP Sant Carles de la Ràpita

			CAP El Temple CAP L'Ametlla de Mar CAP El Perelló CAP Ulldecona CAP La Sènia PCCs in Andalusia: Centro de Salud Aguadulce Sur Centro de Salud Adra Centro de Salud El Ejido Norte Centro de Salud Roquetas Sur Centro de Salud la Mojonera Centro de Salud Puebla de Vícar
region	Region where the PCC is located	Qualitative	Catalonia: Lleida/Tarragona/Terres de l'Ebre/Costa de Ponent Andalusia: Almería
pcc_area	Type of centre	Qualitative	Urban/Rural
medea	Index of socioeconomic deprivation for each of the participating centres	Quantitative	
num_hp	Number of health professionals working at the centre (general practitioners and nurses)	Quantitative	
mig_density	Migration density (%) of the population assigned to each centre	Quantitative	7.0-52.0
mig_assigned	Number of migrants assigned to the PCCs participating in the study	Quantitative	
mig_visit	Number of migrants who visited their assigned PCCs during the intervention	Quantitative	
group	Arm of the intervention	Qualitative	Intervention/Control
Patient			
sex	Sex of the patient	Qualitative	Female/Male
age	Age of the patient	Quantitative	From 14 years old in Andalusia and 15 years old in Catalonia with no age limit in both sites
area_birth	Area of birth of the patient	Qualitative	Africa/Latin-America/Asia/Eastern Europe
date_entry	Date of entry to the PCC	Quantitative	MM/DD/YYYY: Up to 12/31/2024
date_exit	Date of exit to the PCC	Quantitative	MM/DD/YYYY: Up to 12/31/2024
visits	Number of visits to the healthcare centre	Quantitative	

date_visits	Date of visits to the healthcare centre	Quantitative	MM/DD/YYYY: Up to 12/31/2024
lab_hgb	Haemoglobin	Quantitative	
lab_plt	Platelets	Quantitative	
lab_wbc	Leucocytes	Quantitative	
lab_neut	Neutrophils	Quantitative	
lab_lym	Lymphocytes	Quantitative	
lab_eos	Eosinophils	Quantitative	
lab_crea	Creatinine	Quantitative	
lab_urea	Urea	Quantitative	
lab_ast	Aspartate aminotransferase (AST)	Quantitative	
lab_alt	Alanine aminotransferase (ALT)	Quantitative	
lab_ggt	Gamma-glutamyl transpeptidase (GGT)	Quantitative	
lab_bili	Bilirubin	Quantitative	
lab_act	Coagulation	Quantitative	
alert_hiv	Patients fulfilling the criteria for the screening of HIV	Qualitative	Yes/No
alert_syph^	Patients fulfilling the criteria for the screening of syphilis	Qualitative	Yes/No
alert_hbv	Patients fulfilling the criteria for the screening of HBV	Qualitative	Yes/No
alert_hcv	Patients fulfilling the criteria for the screening of HCV	Qualitative	Yes/No
alert_activetb	Patients fulfilling the criteria for the screening of active TB	Qualitative	Yes/No
alert_latenttb^	Patients fulfilling the criteria for the screening of latent TB	Qualitative	Yes/No
alert_chagas	Patients fulfilling the criteria for the screening of *T.cruzi* infection	Qualitative	Yes/No
alert_strongy	Patients fulfilling the criteria for the screening of *S.stercoralis* infection	Qualitative	Yes/No

alert_schisto	Patients fulfilling the criteria for the screening of *Schistosoma spp*. infection	Qualitative	Yes/No
alert_intpar^	Patients fulfilling the criteria for the screening of intestinal parasites	Qualitative	Yes/No
alert_fgm	Patients fulfilling the criteria for the screening of FGM	Qualitative	Yes/No
sero_hiv	Serologies for HIV	Qualitative	Yes/No
sero_syph^	Serologies for syphilis	Qualitative	Yes/No
sero_hbv	Serologies for HBV	Qualitative	Yes/No
sero_hcv	Serologies for HCV	Qualitative	Yes/No
cxray	Chest radiographies for active TB	Qualitative	Yes/No
tst^	TST for latent TB	Qualitative	Yes/No
igra^	IGRA for latent TB	Qualitative	Yes/No
sero_chagas	Serologies for *T.cruzi* infection	Qualitative	Yes/No
sero_strongy	Serologies for *S.stercoralis* infection	Qualitative	Yes/No
sero_schisto	Serologies for *Schistosoma spp.* infection	Qualitative	Yes/No
stool_intpar^	Stool samples for intestinal parasites	Qualitative	Yes/No
fgm	Questionnaires completed for FGM	Qualitative	Yes/No
dx_hiv	Diagnosis of HIV	Qualitative	Yes/No
cd4	CD4 cell count	Quantitative	
hiv_vl	HIV viral load	Quantitative	
dx_syph^	Diagnosis of syphilis	Qualitative	Yes/No
dx_std	Diagnosis of other sexually transmitted diseases (STDs)	Qualitative	Yes/No
dx_hbv	Diagnosis of HBV	Qualitative	Yes/No
dx_hcv	Diagnosis of HCV	Qualitative	Yes/No
dx_activetb	Diagnosis of active TB	Qualitative	Yes/No
dx_latenttb^	Diagnosis of latent TB	Qualitative	Yes/No
dx_chagas	Diagnosis of *T.cruzi* infection	Qualitative	Yes/No
dx_strongy	Diagnosis of *S.stercoralis* infection	Qualitative	Yes/No

dx_schisto	Diagnosis of *Schistosoma spp.* infection	Qualitative	Yes/No
dx_intpar^	Diagnosis of intestinal parasites	Qualitative	Yes/No
dx_fgm	Diagnosis of FGM cases	Qualitative	Yes/No
referral	Referral to specialists due to a diagnosis during the intervention	Qualitative	Yes/No
other_dx	Other diagnoses made during the intervention (besides from the infections included in the study and FGM)	Qualitative	ICD-9 or ICD-10
com_dx	Comorbidities	Qualitative	ICD-9 or ICD-10
immune_dx	Immunosuppressive diagnoses during the intervention period	Qualitative	ICD-9 or ICD-10
drugs	Drugs prescribed to patients during the intervention	Qualitative	Anatomical Therapeutic Chemical (ACT) codes
immune_drugs	Immunosuppressive drugs prescribed during the intervention period	Qualitative	ACT codes
doses	Prescribed drug doses	Quantitative	
tto_start	Treatment start date	Quantitative	MM/DD/YYYY: Up to 12/31/2024
tto_end	Treatment end date	Quantitative	MM/DD/YYYY: Up to 12/31/2024
cost	Cost of the prescribed drugs	Quantitative	In euros (€)
Hospital (MBDS)			
hosp	Hospitalisation	Qualitative	Yes/No
hosp_dx	Primary diagnosis	Qualitative	ICD9 or ICD-10
hosp_dxother	Other diagnosis	Qualitative	ICD9 or ICD-10
hosp_dates	Hospitalisation dates	Quantitative	MM/DD/YYYY: Up to 12/31/2024
hosp_hgb	Haemoglobin	Quantitative	
hosp_plt	Platelets	Quantitative	
hosp_wbc	Leucocytes	Quantitative	
hosp_neut	Neutrophils	Quantitative	
hosp_lym	Lymphocytes	Quantitative	
hosp_eos	Eosinophils	Quantitative	
hosp_crea	Creatinine	Quantitative	

hosp_urea	Urea	Quantitative	
hosp_ast	AST	Quantitative	
hosp_alt	ALT	Quantitative	
hosp_ggt	GGT	Quantitative	
hosp_bili	Bilirubin	Quantitative	
hosp_act	Coagulation	Quantitative	
hosp_sympt	Symptoms	Qualitative	Fever, diarrhoea, rash, fatigue, ulcers, among others
hosp_uci	Need for an Intensive Care Unit	Qualitative	Yes/No
hosp_vent	Need for mechanical ventilation	Qualitative	Yes/No
hosp_drugs	Drugs prescribed	Qualitative	ACT codes
hosp_doses	Drug doses	Quantitative	
hosp_tto_start	Treatment start date	Quantitative	MM/DD/YYYY: Up to 12/31/2024
hosp_tto_end	Treatment end date	Quantitative	MM/DD/YYYY: Up to 12/31/2024
hosp_drugs_cost	Cost of the drugs	Quantitative	In euros (€)
hosp_cost	Cost attributed to hospitalisation	Quantitative	In euros (€)
Hospital follow-up (MBDS)			
followup	Number of follow-up visits after hospitalisation	Quantitative	
followup_dates	Dates of follow-up visits after hospitalisation	Quantitative	MM/DD/YYYY: Up to 12/31/2024
followup_dx	Diagnosis made during the follow-up visits	Qualitative	ICD9 or ICD-10
followup_hgb	Haemoglobin	Quantitative	
followup_plt	Platelets	Quantitative	
followup_wbc	Leucocytes	Quantitative	
followup_neut	Neutrophils	Quantitative	
followup_lym	Lymphocytes	Quantitative	
followup_eos	Eosinophils	Quantitative	
followup_crea	Creatinine	Quantitative	
followup_urea	Urea	Quantitative	

followup_ast	AST	Quantitative	
followup_alt	ALT	Quantitative	
followup_ggt	GGT	Quantitative	
followup_bili	Bilirubin	Quantitative	
followup_act	Coagulation	Quantitative	
followup_drugs	Drugs prescribed during follow-up visits	Qualitative	ACT codes
followup_doses	Drug doses	Quantitative	
followup_tto_start	Treatment start date	Quantitative	MM/DD/YYYY: Up to 12/31/2024
followup_tto_end	Treatment end date	Quantitative	MM/DD/YYYY: Up to 12/31/2024
followup_cost	Cost of the drugs	Quantitative	In euros (€)
Health events			
cured	Cured patients	Qualitative	Yes/No
mortality	Patients who died during the intervention period	Qualitative	Yes/No

PC: Primary care, PCT: Primary care team, EAP: Equipo de Atención Primaria, PCC: Primary care centre, CAP: Centro de Atención Primaria, HIV: human immunodeficiency virus, HBV: hepatitis B virus, HCV: hepatitis C virus, TB: tuberculosis, FGM: female genital mutilation, ACT: Anatomical Therapeutic Chemical, ICD-9: International Classification of Diseases Ninth Revision, ICD-10: International Classification of Diseases Tenth Revision, MBDS: Minimum basic dataset; ^Only in the Andalusian site