

INPLASY202660102

doi: 10.37766/inplasy2026.6.0102

Received: 21 June 2026

Published: 22 June 2026

Stocco, C; Borges, PKO; Muller, EV.

Corresponding author:

Caroliny Stocco

carolinystocco@hotmail.com

Author Affiliation:

State University of Ponta Grossa (UEPG), Brazil.

ADMINISTRATIVE INFORMATION

Support - Does not apply.

Review Stage at time of this submission - Preliminary searches.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202660102

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 22 June 2026 and was last updated on 22 June 2026.

INTRODUCTION

Review question / Objective **R** question: How has probabilistic linkage been applied to the integration of health information systems, with emphasis on immunization information systems and public health surveillance?

Objective: To map and synthesize the scientific evidence on the use of probabilistic linkage for integrating health information systems, identifying its applications, methodological approaches, validation strategies, and contributions to data quality and public health surveillance, with emphasis on immunization information systems.

Background Health information systems are essential tools for planning, monitoring, and evaluating public health actions. In immunization programs, these systems support vaccination coverage monitoring, identification of susceptible populations, and epidemiological surveillance activities. However, fragmentation of databases,

inconsistencies in data quality, and the absence of unique identifiers often limit interoperability and the effective use of information. Probabilistic linkage has emerged as an important methodological strategy to integrate records from different databases and improve data quality in health research and surveillance.

Rationale The growing use of probabilistic linkage in health information systems has generated a broad and heterogeneous body of literature. Although several studies have explored record linkage methods in health databases, evidence regarding their application in immunization information systems and public health surveillance remains scattered. Mapping this evidence is essential to identify methodological approaches, validation strategies, implementation challenges, and knowledge gaps. The findings may support future research and contribute to the development of data integration methodologies aimed at improving surveillance activities and decision-making in public health.

METHODS

Strategy of data synthesis The review will be conducted in the following electronic databases: PubMed/MEDLINE, Scopus, Web of Science Core Collection, LILACS (via Virtual Health Library), and SciELO.

Search strategies will combine controlled vocabulary (Medical Subject Headings - MeSH and Health Sciences Descriptors - DeCS) and free-text terms related to probabilistic linkage, record linkage, immunization, vaccination, public health surveillance, health information systems, and data quality. The search strategies will be adapted to the indexing system and search syntax of each database to ensure comprehensive retrieval of studies. In addition to identifying studies specifically addressing immunization information systems, the search strategies are intended to capture methodological studies on probabilistic linkage applied to other health information systems whenever they provide evidence relevant to the integration and qualification of health data.

PubMed/MEDLINE
 ("Record Linkage"[Mesh] OR "record linkage" OR "probabilistic linkage")
 AND
 ("Health Information Systems" OR "Electronic Health Records")
 AND
 (Immunization OR Vaccination OR "Immunization Programs" OR "Public Health Surveillance")

Scopus
 TITLE-ABS-KEY ("record linkage" OR "probabilistic linkage")
 AND TITLE-ABS-KEY ("health information systems" OR "electronic health records")
 AND TITLE-ABS-KEY (immunization OR vaccination OR "immunization programs" OR "public health surveillance")

Web of Science Core Collection
 TS=("record linkage" OR "probabilistic linkage")
 AND TS=("health information systems" OR "electronic health records")
 AND TS=(immunization OR vaccination OR "immunization programs" OR "public health surveillance")

LILACS (BVS)
 ("Relacionamento de Registros")
 AND
 ("Sistemas de Informação em Saúde")
 AND
 ("Imunização" OR "Vacinação" OR "Vigilância em Saúde")

SciELO
 ("record linkage" OR "probabilistic linkage" OR "relacionamento de registros")
 AND
 ("health information systems" OR "sistemas de informação em saúde")
 AND
 (immunization OR vaccination OR imunização OR vacinação OR "public health surveillance" OR "vigilância em saúde")

The reference lists of all included studies will also be manually screened to identify additional eligible studies that may not have been retrieved through the electronic database searches.

Retrieved studies will be synthesized descriptively and narratively according to the objectives of the review. Evidence will be grouped by application area, database type, probabilistic linkage method, software used, validation metrics, and contribution to health information systems and public health surveillance.

Eligibility criteria

Population:

Studies involving health databases, immunization records, vaccination records, or health information systems.

Concept:

Probabilistic linkage or record linkage methods used to integrate records from different databases.

Context:

Health information systems, public health surveillance systems, immunization information systems, epidemiological databases, and related health data environments.

Types of evidence sources:

Original studies, methodological studies, validation studies, implementation studies, observational studies, and mixed-methods studies. Only studies published in indexed scientific journals will be considered.

Languages:

English, Portuguese, and Spanish.

Publication period:

No publication date restrictions will be applied.

Source of evidence screening and selection All records identified through database searches will be exported to a Microsoft Excel spreadsheet, where duplicate records will be identified and removed. The screening process will be conducted by the principal reviewer through sequential evaluation of titles, abstracts, and full-text articles according to the predefined eligibility criteria. Potentially eligible studies will undergo full-text assessment for final inclusion. Reasons for exclusion at the full-text stage will be documented

whenever applicable. Methodological supervision and verification of critical stages of the study selection process will be performed by the supervisor and co-supervisor, including verification of eligibility decisions, to ensure consistency and methodological rigor. The study selection process will be reported using a PRISMA-ScR flow diagram.

Data management All references retrieved from the electronic databases will be organized in Microsoft Excel. Duplicate records will be identified and removed manually. Data from the included studies will be extracted by the principal reviewer using a standardized data extraction form developed specifically for this review. Methodological verification of data extraction and synthesis will be carried out by the supervisor and co-supervisor at critical stages of the review process to ensure accuracy and consistency. The extraction spreadsheet will include study characteristics, database type, probabilistic linkage methods, matching variables, software used, validation metrics, and main findings. The data extraction form may be refined during the review process if necessary.

Reporting results / Analysis of the evidence A descriptive and narrative synthesis will be conducted. Evidence will be grouped according to the area of application, type of database, linkage method, software used, validation strategy, and contribution to data quality and public health surveillance. Findings will be presented in tables, figures, and thematic categories.

Presentation of the results The results will be presented using a PRISMA-ScR flow diagram, summary tables, and thematic evidence maps. Tables will summarize study characteristics, methodological approaches, software tools, validation metrics, and applications of probabilistic linkage. Evidence gaps and future research opportunities will also be highlighted. The presentation of results will follow the PRISMA-ScR reporting guideline.

Language restriction Searches will be restricted to studies published in English, Portuguese, and Spanish.

Country(ies) involved Brazil.

Other relevant information This scoping review is part of a doctoral research project focused on the integration of immunization information systems and the development of probabilistic linkage

methodologies to improve public health surveillance.

Keywords Immunization; Health Information Systems; Record Linkage; Data Quality; Public Health Surveillance.

Dissemination plans The results of this review will be disseminated through publication in a peer-reviewed scientific journal, presentation at scientific conferences, and incorporation into a doctoral thesis. Findings will also support the development of methodological approaches for integrating immunization information systems in public health settings.

Contributions of each author

Author 1 - Carolyn Stocco - Conceptualization, methodology, literature search, study selection, data extraction, data synthesis, writing of the original draft, and project administration.

Email: carolynstocco@hotmail.com

Author 2 - Pollyanna Kassia de Oliveira Borges - Methodological supervision, critical review of study selection and data extraction, interpretation of findings, manuscript review and editing, and supervision.

Email: pkoborges@uepg.br

Author 3 - Erildo Vicente Muller - Methodological co-supervision, critical review of data synthesis, interpretation of findings, manuscript review and editing.

Email: erildomuller@hotmail.com