Journal of Medical Systems

Barriers, limitations, and challenges to the implementation of artificial intelligence in Latin American healthcare systems: a scoping review

Valentina Velasco-Muñoz, Pablo Pineda-Sanabria, Juan Diego Navarro-Sánchez, Daniela Varón-Jauregui, Lucas Gallego-Tristancho, Mateo Uribe-Gaviria, Santiago Andrés Suárez-Gómez, Luis M. Ahumada

Corresponding author: Valentina Velasco-Muñoz, Department of Internal Medicine, Pontificia Universidad Javeriana, Cra 7 No 40 - 62 , Bogotá D.C., Colombia. Email: velascomv@javeriana.edu.co; telephone: +57-315-632-0743.

Online Resource 4. Coding manual and coding appendix

Supplement to: Barriers, limitations, and challenges to the implementation of artificial intelligence in Latin American healthcare systems: a scoping review.

This resource has two parts. **Part A** reproduces the coding rules of the manual under which every included source was recharted (version 1.10, frozen 13 September 2026, before recharting began). Sections 3 to 11 of the manual are reproduced as frozen, except for two sentences of section 11 addressed to the review team rather than stating a coding rule; sections 1, 2 and 12 (the rationale for recharting, the division of labour, and the change log) are working-process text and are included in the Open Science Framework deposit with the full manual. **Part B** is the coding appendix: the dated applications of those rules recorded while charting, the controlled vocabulary as it stood when charting closed, the locator correction, the material the extraction could not read, the verification record, and the reconciliation of the first-round instances against the recharted dataset. Nothing in Part B amends the manual; each entry records how a frozen rule was applied to a case the manual had not spelled out, with the rows it governed.

Row identifiers (for example, Machado R53) refer to the deposited charting dataset, in which every row carries its verbatim quotation, page locator and coder note.

Part A. Coding manual, version 1.10 (frozen 13 September 2026), sections 3 to 11

Sections 3 to 11 are reproduced as frozen, except that two sentences of section 11 addressed to the review team rather than stating a coding rule are omitted. The complete manual, including sections 1, 2 and 12, is in the deposited record.

3. The unit of analysis

The unit is the **barrier instance**: one distinct claim, made in one source, that something impedes the implementation or adoption of AI or of the digital infrastructure AI depends on, in a Latin American healthcare setting.

A barrier instance is charted when all of the following hold.

1. The source states it as an impediment, a limitation, a challenge, a gap, a risk or an unmet precondition. Recommendations count only where they name the deficit they address, see section 4.
2. It is attributable to a quotable span of text in the source.
3. It is distinct from every claim already charted for that source, in the sense set out in section 4.
4. **It is not attributed specifically to a setting outside Latin America and the Caribbean.**
5. **What it impedes is AI or the digital infrastructure AI depends on**, not the source's clinical or research field in general. See the concept exclusion below.

The concept exclusion rule

Added 12 September 2026 on Pablo's decision, after Werutsky surfaced it. Sources often discuss barriers to their own field alongside barriers to AI and digital health. A Lancet Oncology Series paper on cancer control in the region charts, in the original 829, claims such as the region spending 0.7% of GDP on research and development, regulations requiring sponsors to cover standard-of-care costs, and patenting bureaucracy defeating technology transfer. Each is a real barrier and none impedes an AI application or digital infrastructure: they impede clinical research.

A claim is charted only when what it impedes is an AI application, or digital infrastructure, data or systems. A claim about the source's field at large, cancer research funding, trial regulation, workforce shortages in oncology, is not charted, and ai_or_digital has no honest value for it.

The exception, and it is charted with its reason recorded. Where the source itself ties the deficit to AI, big data, digital health or personalised medicine, the claim is in concept and the linking words go in coder_note so the decision can be defended row by row. Default to excluding; keep only what the source links itself. Do not supply the link on the source's behalf.

Excluded claims go on the crosswalk sheet with disposition out_of_concept. Expect this to bite hardest in Series papers, national cancer or health plans, and reviews whose subject is a clinical field rather than technology.

This tightens the corpus in the same direction as the regional rule, and it answers the editor's third objection at the point where it actually bites: a reviewer asking what counts as an AI barrier is asking this question.

The regional exclusion rule, instance level

Added 12 September 2026 on Pablo's instruction, after the first verified batch. Eligibility had been applied only at source level: a source entered if it reported data specific to at least one Latin American country. From multinational sources that admitted individual claims the source itself attributes to Africa or Asia, which do not meet the review's own criterion.

A claim is charted when its setting is Latin American, regional, unspecified, or a wider grouping the region belongs to. A claim attributed specifically to a non-regional country or region is not charted.

Setting of the claim	Charted
A named Latin American or Caribbean country	yes
Latin America or the Caribbean as a region	yes
A wider grouping such as LMICs or the global south, pooled	yes
No geographic referent at all, a general claim	yes
Attributed specifically to a non-regional country or region, "In Africa...", "In Asia..."	**no**
Referent undeterminable	yes, flagged unclear

The discriminator is what the **claim** is attributed to, not what illustrates it. A general finding about regulatory capacity in LMICs, illustrated by a Kenyan participant's quote, is a general claim and is charted: the claim is not about Kenya, the illustration is. A statement beginning "In Africa the existence of multiple local languages makes AI implementation difficult" is about Africa and is not charted.

Excluded claims go on the crosswalk sheet with disposition other_region_specific rather than being silently dropped, so the count stays reportable. This **tightens** eligibility fidelity and belongs in the Methods: it is the instance-level counterpart of the source-level regional criterion and answers audit item P2-2 more directly than the source-level analysis alone. Two Baka instances, rows 42 and 43, are excluded under it.

Compound statements are split by sub-theme. A sentence naming three different deficits produces three rows. This is the convention the original 41 were charted under and decision D3 kept it, so it does not change here. A sentence naming one deficit affecting three groups produces one row, because the deficit is the unit, not the affected party.

4. chart_status: CORE or RESTATEMENT

This is the rule the rechart exists for, so it is written to be applied mechanically.

Definition. A statement is a RESTATEMENT when it re-expresses a claim already charted for the same source, adding no new referent, no new mechanism and no new evidence. It is CORE otherwise.

The test, applied in order.

1. Is there an existing CORE row for this source with the same sub_theme? If no, the statement is CORE.
2. If yes, does the new statement differ from that row in **any** of: the actor or institution it concerns; the technology or system it concerns; the setting, level of care or population it concerns; the causal mechanism it names; or the specific content of the deficit? If it differs in any of these, it is CORE.
3. If it differs in none of them, it is a RESTATEMENT, no matter where it appears, how it is worded, or how much rhetorical weight it carries.

Consequences of the test, spelled out so they are not re-litigated per source.

- A recommendation that mirrors a barrier already charted is a RESTATEMENT. "Training programmes should be expanded" after "staff lack training" adds nothing. If the recommendation names a deficit not charted anywhere in that source, it is CORE.
- Abstract, Discussion and Conclusions restatements of a Results finding are RESTATEMENTS. This is where most of them live.
- A repeated key is **not** by itself a restatement. The corpus deliberately permits several instances on one level, domain and sub-theme key when the claims are separately stated, which is exactly why decision D3 kept the four Portilho equity facets separate: they sit under three different domains and are separately measured. Step 2 of the test, not the key, decides.
- **Where the same claim appears once without data and once with data, chart one CORE row**, place it at the data-bearing statement, quote that statement, and code evidence_type from it. Recording the measured version as a restatement of the unmeasured one would discard the evidence and bias the headline.
- Where two statements of one claim differ only in specificity, the more specific one is the CORE row.

Every restatement is charted as its own row, carrying chart_status = RESTATEMENT and restatement_of set to the barrier_id of its CORE row. Restatements are excluded from every count. They are kept rather than dropped so the decision stays auditable and a reviewer who disagrees can recompute without re-reading anything.

5. evidence_type: empirically studied or asserted

Frozen under decision D5, scenario B. The operative rule, which is also the Methods sentence that replaces the current one:

> A barrier was classified as **empirically studied** when the source reported data collected for the purpose of characterising that barrier, whether quantitative (survey frequencies, measured indicators, performance or utilisation data) or qualitative (interviews, focus groups, structured stakeholder consultation, or documented content or thematic analysis of primary material), and when the barrier statement was traceable to those data rather than to the authors' commentary. Barriers stated without reference to data collected in the source, including barriers cited from other publications without primary examination, were classified as **asserted**.

Applying it:

- **A review that runs its own documented analysis is reporting its own data.** Arcos-Huaman's "Regulaciones (31%)" comes from that review's own scientometric analysis and is empirically studied. Werutsky's "0.7% of gross domestic product", cited to its reference 58, is someone else's statistic and is asserted. 120 rows in the original 829 sit in review, perspective or document-analysis sources and are currently coded empirically studied; both cases hide in that number, and separating them is expected to be the single largest driver of the change in the headline, larger than the restatement removal.

- The test is about **this source's own data**, not about the source's design. A primary study's Discussion paragraph that speculates beyond its data is asserted. A narrative review that reports its own documented content analysis of policy texts is empirically studied for the claims that analysis supports.
- A barrier attributed to another publication is asserted, unless this source re-examined the primary material and says so. Section 5A records *whose* claim it is, which is a separate question from whether it is supported.
- Numbers are not required. A thematic analysis of interviews that identifies a barrier as a theme is empirically studied. A count is sufficient but not necessary.
- Where the source reports data but the specific barrier statement is not traceable to them, it is asserted. Traceability, not proximity, is the test.
- **Scenario C is not coded and is not reported.** No third column, no sensitivity analysis on a broader rule. The disposition to the editor is that the rule was widened to count primary qualitative research, not that two rules are on offer.
- evidence_type_pre_recode carries the original coding for the 829 original instances, never overwritten, so the point-by-point can state exactly how many instances moved and in which direction.

5.1 The Abstract carries no provenance information of its own

Added at v1.10 on Pablo's annotation, made while verifying Bellei. Citation markers are suppressed in abstracts by journal convention, so their **absence there is a formatting fact, not evidence that the claim is the source's own**. A coder who reads for citation markers will default an entire Abstract or Highlights block to primary and empirically studied, which is the proximity test creeping back in through a side door after section 5 expelled it. The same applies to any front-matter summary block: Abstract, Highlights, Key Points, Graphical Abstract, Plain Language Summary, Lay Summary.

Procedure. For any statement located in such a block, find its counterpart in the body first.

1. **A counterpart exists in the body.** The body statement is the CORE row. Quote it from the body, code it from the body, and locate it at the body page. The front-matter statement is a RESTATEMENT of that row under section 4. No further judgement is needed, because evidence_type on a restatement never enters a count.
2. **No counterpart exists anywhere in the body.** The claim lives only in the front matter, so it is CORE, and it defaults to asserted and unattributed. A claim with no body text behind it has nothing traceable behind it.
3. **Exception, narrow.** Code empirically studied and primary only where the front-matter sentence reports a result the source's own Methods produced **and** the body carries the underlying data: a table, a figure, a count, or a stated denominator, even where no body sentence restates the claim in words. Name the table, figure or section in coder_note.

A CORE row's page_locator never points at the front matter while the body carries the same claim.

This specialises the section 4 rule that one CORE row is placed at the data-bearing statement, and the section 5 rule that traceability rather than proximity is the test. Abstract-against-body is the commonest instance of both, and neither named it. Nothing here changes what counts as data; it changes only where the coder is required to look before deciding.

5A. claim_provenance: whose observation is this

The evidence rule says whether a claim is supported. This says where it came from. The two are independent: a review can cite another paper's measured finding, which is asserted under section 5 and cited_external here.

Value	When
primary	The source reports its own data, its own documented analysis, or its own first-hand institutional experience.
cited_internal	The source attributes the claim to a publication that is itself one of the 43 included sources.
cited_external	The source attributes the claim to a publication outside the corpus.
unattributed	The source asserts the claim with no citation and no data of its own.

How cited_internal is determined. Not by textual matching: no two sources in the corpus share a verbatim, because a source citing another paraphrases it. Resolve the citation marker against that source's own reference list, then check that reference against the list of 43 included sources. This is done here and not in NotebookLM, which sees only the sources loaded into a given notebook and will therefore answer this column unreliably; its cited_internal answers are treated as unverified.

claim_source_ref carries the citation as printed, verbatim, whenever the value is cited_internal or cited_external: the bracketed number or author-year marker, plus the first author and year from the source's own reference list where they can be read off without effort. This is what makes the citation chain visible. Recording only the category tells us how much of the corpus is recycled; recording the reference tells us which findings are being recycled and how often, which is the difference between conceding non-independence and measuring it.

Why this is charted rather than inferred. Twelve systematic or scoping reviews and twelve narrative reviews sit in the corpus and cite overlapping literature, so one primary finding can enter as four instances through four reviews. The restatement rule at section 4 catches repetition inside one source and is blind to repetition across sources. This column is the instrument that catches the second kind, and it is the direct answer to the editor's objection that the analysis "mixes primary studies, reviews, policy analyses, and opinion pieces while allowing multiple counts from individual publications".

A cited claim is still charted. A review that cites a study from outside the region and then asserts the barrier applies regionally is making a claim about the region, and the corpus is a map of what the regional literature reports. Chart it, code the provenance, and let the analysis separate the layers. Do not silently drop it.

5A.1 The redundancy analysis, and why cited claims are not excluded

Specified 11 September 2026 after Pablo raised whether cited claims double-count.

Cited claims are charted and counted. Three reasons, recorded so this is not reopened.

1. **Excluding them would assume the conclusion.** The paper's headline is that most barrier claims in this literature are asserted rather than measured, and a claim attributed to another publication is asserted by definition. Deleting them would collapse the asserted share because the asserted instances were removed, not because the literature improved. The finding would become an artefact of the filter.
2. **It would change the review's scope.** Twenty-four of the 43 sources are reviews. Excluding cited claims removes most of what they contribute and converts a map of what the regional literature reports into a map of what primary regional research observed, which is a different review and contradicts the registered scope.
3. **The independence concern is already answered by the study-level denominator.** If four reviews cite the same primary study, the instance count says four and the study-level count says four of 43 sources report that barrier, which is the honest statement about a literature. Propagation across four publications is a real signal about the discourse and the study-level figure is what stops a reader taking it for four independent observations.

The redundancy analysis, reported as a supplementary table. This measures the dependence structure rather than conceding it in a Limitations sentence, and it costs nothing additional because claim_source_ref is charted anyway.

- The count of cited_internal instances, with the corpus sources they trace back to, and a **deduplicated instance count** reported alongside the primary count: the primary count as charted, and the count after collapsing instances that trace to a claim already charted from its own source.
- For cited_external, the concentration of citation: how many instances across how many sources descend from how many external publications, with the most-cited works named.
- The same figures broken down by study design, since the expectation is that narrative reviews carry most of the redundancy and primary studies almost none. If that holds it strengthens the design-stratified analysis at audit item P1-4; if it does not, it is reported as it falls.

The deduplicated count is a **secondary** figure and does not replace the headline. It is what allows the point-by-point to answer the editor's non-independence objection with a measurement rather than an acknowledgement.

5B. ai_or_digital: what the barrier is a barrier to

Coded per instance, not per source. The source-level flag at section 7 answers the editor's third objection weakly, because 39 of 41 sources name an AI technology somewhere, which says almost nothing about what any individual barrier statement concerns.

Value	When
ai_specific	The barrier concerns an AI application: machine learning, deep learning, clinical decision support, natural language processing, computer vision or generative AI.
digital_general	The barrier concerns digital infrastructure, connectivity, interoperability, health information systems or digital capacity, without reference to an AI application.
both	The statement explicitly spans the two, for example a data-quality deficit stated as blocking both record-keeping and model training.

The discriminator is what the barrier is a barrier *to*, not the nature of the deficit. The deficit type is already carried by domain, so do not code an infrastructural deficit as digital_general merely because it is infrastructural. Ask what the statement says is impeded. "No national interoperability standard exists" in a paper on AI readiness is digital_general if the statement concerns health information systems, and ai_specific if it names model deployment or algorithmic decision support as the thing blocked. Where the source explicitly frames the infrastructure deficit as a precondition for AI, that is both. Without this rule every infrastructural barrier drifts into whichever bucket the coder had in mind that hour, and the resulting split means nothing.

Code from the statement, not from the source's title or aims. A paper about AI readiness that states "connectivity in rural facilities is intermittent", with no AI referent in the statement or its immediate context, is making a digital_general claim.

How the distinction is reported

Decided 11 September 2026, revised the same day after Pablo's proposal, which is the version that stands. **One corpus, one headline denominator, the distinction carried as a coded flag and reported as a result.**

1. **Charting scope is unchanged and the corpus is not split.** Every barrier statement in the 43 included sources is charted whatever it impedes, and all of them sit in one analytical set. Eligibility admitted both concepts and was registered that way, so there is no post hoc narrowing to defend.
2. **ai_or_digital is coded on every instance** and reported, not used to filter. The Results give the proportion in each of the three values, overall and by level and domain, so a reader can see where the two concepts concentrate.
3. **The sensitivity analysis runs on the ai_specific plus both subset**, reporting whether the level and domain distribution and the direction of the empirically studied against asserted finding hold when only AI-specific barriers are counted. If they do not hold, that is a finding and it is reported as one, not softened.
4. **Why this rather than making the AI subset primary.** Three reasons, recorded so the choice is not revisited on instinct. The headline then moves only once, for the rechart, instead of twice, which matters in front of an editor who found the previous analysis hard to interpret. It answers what the letter actually asked, which is to clarify the distinction, not to restrict the analysis. And it protects the subgroup analyses: the sub-theme rates, the level-by-domain map and the design stratification sit on thin cells already, and a narrowed denominator would make several of them uninterpretable.
5. **The title stands**, and so does the Abstract, subject to one condition. If the ai_specific plus both share comes back below roughly half of instances, describing the corpus as barriers to AI implementation without qualification is strained, and the Abstract and Methods gain one clause saying the review covers barriers to AI and to the digital infrastructure it depends on. That is decided when the number exists, never before, and the Methods rationale for admitting both concepts is already drafted and already sound: infrastructure and interoperability are preconditions for deployment rather than a separate subject.
6. **The source-level flag at section 7 is retained for Table 1.** The P1-7 replacement text on file, stating that 39 of 41 sources name an AI technology and that the two that do not contribute 32 of 829 instances (3.9%),

describes that source-level flag only. **It is superseded and must be rewritten at step 4.3** against the instance-level numbers, which are the ones that answer the editor.

5C. latam_attributable: is this claim about the region

Coded per instance. The regional_specificity variable at section 7 is source-level, so for the eight multinational sources contributing 206 instances it says only that the source is mixed. This says whether a given claim is attributable to the region.

Value	When
named_country	The statement concerns one or more named Latin American or Caribbean countries. Record which in coder_note.
regional	The statement concerns Latin America or the Caribbean as a region, without naming a country.
pooled_multiregion	The statement concerns a wider set (LMICs, the global south, a multi-region sample) that the region belongs to, with no separately attributable regional content. Charted.
other_region_specific	The statement is attributed specifically to a country or region outside Latin America and the Caribbean. **Not charted**, recorded on the crosswalk instead, per the exclusion rule at section 3.
unspecified_setting	The statement makes a general claim about AI or digital health with no geographic referent at all, typically in an Introduction framing passage. The referent is clear, it is simply not geographic.
unclear	The referent cannot be determined from the statement and its immediate context. Use sparingly.

pooled_or_external, used up to v1.6, is split by v1.7 into pooled_multiregion and other_region_specific, because the first is charted and the second is not. Rows coded under the old value are recoded.

Sources coded latam_exclusive at source level will be named_country, regional or unspecified_setting throughout, so this costs little on the single-region sources and does its real work on the multinational ones.

unspecified_setting is analytically useful rather than a dustbin: it separates the barriers the regional literature reports about the region from the generic framing claims it imports from elsewhere, which is a distinction a reader of a regional scoping review will want and which no existing variable captures.

5D. Crosswalk to the original charting

The corpus is being recharted, not rebuilt, and the editor will ask what happened to the 829. Every count we report has to be traceable back to what it replaced.

- Each new row carries original_barrier_id, the barrier_id of the instance it corresponds to in C_barriers_included, or blank where the row is genuinely new. Where a new CORE row absorbs several original rows, list them separated by semicolons.
- Each original instance with no counterpart is recorded on the crosswalk sheet with one disposition: restatement (now flagged and excluded from counts), merged (absorbed into another row, which is named), not_a_barrier (fails section 3), out_of_concept (impedes the source's field rather than AI or digital infrastructure), superseded (replaced by a data-bearing statement of the same claim), or retained_unchanged.

This is cheap while the two tables are side by side and unreconstructable afterwards. Without it the rechart reads as a new dataset rather than a correction, which is a much weaker position in front of an editor who has already declined the paper once.

5E. barrier_status: experienced or anticipated

Coded per instance. The distinction is between a barrier that arose in an implementation the source describes and a barrier projected for implementations not yet undertaken.

Value	When
experienced	The source describes a deployed or piloted system, an existing programme or an actual institutional experience, and the barrier arose within it.
anticipated	The barrier is projected, expected, feared or reasoned about in advance of an implementation the source describes.
both	The statement covers an observed difficulty and its projected persistence or escalation.
unclear	The statement gives no basis for deciding. Use sparingly; an overused unclear destroys the variable.

This is independent of evidence_type and of claim_provenance, and the three must not be collapsed. A survey can measure anticipated concerns, which is empirically studied and anticipated. A narrative review can assert a difficulty observed in a deployment it cites, which is asserted, cited_external and experienced.

Why it is worth the extra judgement on roughly 900 rows: the regional literature is largely anticipatory, and if that is true it explains the asserted-not-measured finding rather than merely restating it. A barrier nobody has yet met cannot have been measured. That is a mechanism, and it is a stronger paper than one reporting the same proportion with no account of why.

6. Controlled vocabulary

Fixed. Values outside these lists are not coded without an entry in section 12.

level, five values: health-system, regulatory/political, institutional, professional, patient.

domain, seven values: regulatory, operational, financial, infrastructural, sociocultural, ethical, organisational.

sub_theme, 25 values, as carried by the original corpus: governance/regulatory framework gap; trust/acceptance; equity/access; data quality/availability; connectivity/infrastructure gap; cost/procurement/funding scarcity; digital literacy/workforce capacity gap; workflow integration; data protection/privacy; poor interoperability; workforce shortage/multidisciplinary gap; liability/ethical risk; model validation/generalizability gap; weak collaborative research networks; technological dependency/sovereignty; algorithmic bias/fairness; regulatory heterogeneity/harmonization gap; change management/stakeholder engagement; clinical validation gap; environmental/energy sustainability; explainability/black-box; device/technology approval process; language/cultural adaptation gap; private-sector influence on governance; bureaucratic/administrative barriers.

sub_theme_raw keeps the source's own wording where it differs usefully from the controlled term.

Proposing a new sub-theme. Do not invent one silently. Chart the row with the closest existing term, set coder_note to NEW SUBTHEME PROPOSED: <term> with the reason, and Pablo decides. A new term admitted mid-rechart requires a sweep of every source already coded, which is why the bar is high. The Baka charting of 7 September already flagged eight rows as NOVEL PAIRING, meaning a level, domain and sub-theme triple with no precedent in the original corpus; those carry forward and need the same adjudication.

7. Source-level variables, one value per source

regional_specificity, created by decision D1. Three levels, coded from the full text and never from the country column.

Code	Definition
latam_exclusive	All reported data are drawn from Latin American or Caribbean settings, whether one country or several.
latam_arm	The source includes non-regional settings, and Latin American data are separately reported as a

Code	Definition
	study arm, a named country case study, or another separately attributable unit.
latam_incidental	The source includes non-regional settings, and Latin American countries appear only as participants, illustrative cases or members of a larger set, with no separately reported Latin American result.

The discriminator between latam_arm and latam_incidental is **separate attributability of the result**, not the presence of a country name. A source naming Peru in its participant table but reporting only pooled findings is incidental. Multi-country sources drawn entirely from the region are latam_exclusive, not latam_arm: the level is about the absence of non-regional material, not about being single-country, and it must not be conflated with the single-country against multi-country cut in Figure 2b. Reviews and document analyses are coded on the provenance of the material they synthesise, by the same rule. Provisional readings for nine sources are tabulated in 00_START_HERE/01_DO_THIS_NEXT/D1_DECISION_regional_fidelity.md section 2 and are explicitly not verified codings. Baka is contested there, and whatever rule disposes of it must also dispose of McLennan and Marco-Ruiz, which have the same shape.

ai_specific_technology, yes or no, from audit item P1-7. Yes when the source names at least one of machine learning, deep learning, clinical decision support, natural language processing, computer vision or generative AI. No when it addresses digital transformation without reference to any AI application. On the original corpus this was 39 of 41 yes, with the two no sources contributing 32 of 829 instances.

underlying_dataset, free text. The primary data collection the source draws on, where it is not the source's own: a named national survey, cohort, registry or administrative dataset, with its year. Portilho is a secondary analysis of the Cetic.br ICT in Health 2024 survey, and if another included source analyses the same survey, or if a conference version and a full paper both sit in the corpus, that is double counting at source level and it is the same independence objection one layer up. Blank where the source collected its own data or synthesises literature. Detectable now, needs 43 re-reads later.

study_design, assigned under the rule from audit item P2-1: where a source describes its design in compound terms, the design recorded is **the one that governed how its barrier statements were generated**, with the secondary component noted in coder_note. The stratified analysis follows this assignment, not a keyword grouping of the free-text field, which is what made the Table 1 and dataset counts diverge.

8. Locators

page_locator: page number as printed in the source, or the PDF page when the source carries none, labelled p. or pdf p. accordingly. A range where the claim spans pages.

section: the source's own heading, verbatim and untranslated. Use the heading, not a description of it.

Claims stated in a figure, table or box. Chartable when the figure, table or box states the item as a barrier. page_locator names the figure, table or box rather than a page alone ("figure 2", "table 1", "panel: Key priorities"), barrier_verbatim carries the cell text exactly as printed, and coder_note carries the panel or column heading wherever the cell alone is not self-explanatory. 18 rows in the original 829 are fragments of this kind, some only three or four words, and without the heading they cannot be interpreted by a verifier or a reviewer.

The corpus is trilingual: 44 sources in English, 8 in Portuguese, 7 in Spanish, some with multilingual abstracts. Quotations stay in the source's language, never translated, and any tool that pattern-matches this corpus must carry all three languages. A first pass of scripts/sweep_corpus.py with English-only patterns misreported 312 rows as having no AI term, against a true figure of 211, because it did not recognise "IA" or "inteligencia artificial".

barrier_verbatim: the claim in the source's own words, quoted exactly. **Checked mechanically by scripts/validate_rechart.py before any batch is handed over; a batch containing a row that is not an exact substring of its source is not handed over.** The 7 September Baka charting failed this on 38 of 43 rows, which were condensations carrying participant quote codes rather than quotations. Quote first, code second, never the reverse, in the source's own language, not translated and not paraphrased. Trim to the span that carries the claim. Use square brackets for any insertion needed to make the quotation parse, and an ellipsis for any elision.

9. Output schema

One row per barrier instance, in charting_data_v2.xlsx, sheet barriers:

record_id, barrier_id, original_barrier_id, tranche, barrier_verbatim, level, domain, sub_theme, sub_theme_raw, evidence_type, evidence_type_pre_recode, chart_status, restatement_of, claim_provenance, claim_source_ref, ai_or_digital, latam_attributable, barrier_status, page_locator, section, coder_note, Verified (Y/N), first_author, year.

assigned_reviewer is dropped as of v1.7: Pablo verifies every row, so the column carried one value and no information. The fact is recorded here and on the README sheet, which is where a reviewer looks for it.

Sheet sources, one row per included source: the existing A_sources columns plus regional_specificity, ai_specific_technology, underlying_dataset and the reconciled study_design.

Sheet crosswalk, one row per original instance with no counterpart in the rechart, carrying its barrier_id, its disposition under section 5D and, where the disposition is merged, the barrier_id that absorbed it.

tranche takes original_2026-06 or repair_2026-09 so the repair's contribution stays reportable without re-deriving it. A README sheet states which columns apply to which tranche, so a reviewer reads an empty cell as out of scope rather than as missing data.

10. Working order

Sources are recharted in batches of five, in alphabetical order by record_id, with two exceptions: the repair tranche goes first, because those 66 rows are already charted on this schema and verifying them calibrates the rule before it is applied to the other 41; and the five sources with the largest number of repeated keys, Werutsky, Dantas, Bellei, Shahid and Ciecierski-Holmes, go second, because they carry a third of the restatement exposure and will surface any weakness in section 4 early.

No batch starts before the previous one is verified. Nothing downstream of the corpus, meaning any count, table, figure or sentence, moves until every batch is closed.

11. Verification

Pablo is the sole verifier for the rechart and verifies every row against the full text. Verification means each of: the quotation is accurate and present at the stated locator; the claim is a barrier as defined in section 3; chart_status is right under the section 4 test; evidence_type is right under the section 5 rule; claim_provenance, ai_or_digital, latam_attributable and barrier_status are right under sections 5A to 5C and 5E; the crosswalk entry is right under section 5D; and the vocabulary assignment is defensible.

A changed field is recorded in coder_note as CHANGED <field>: <from> -> <to>, so the rate at which verification changed the initial coding is measurable. Rows are marked Verified = Y only when every field has been checked.

'

Part B. Coding appendix

B1. Applications of the manual recorded during charting, by date

Restatement test (manual section 4, step 2), 13 September 2026. A second statement of a deficit by the same source is a distinct instance when it adds a referent named in the statement: a different actor, technology class, setting or mechanism. It is a restatement when the difference is rhetorical position, wording, a consequence clause, or a narrowing of the affected group. Where two statements differ only in specificity, the more specific one is the instance and the other its restatement.

Front matter (manual section 5.1), 13 September 2026. A statement in an abstract, highlights box or key-points box is resolved against its counterpart in the body before it is coded; the instance is located in the body wherever the

body carries the same claim. With no body counterpart, the statement is charted as an instance, asserted and unattributed, because journals suppress citation markers in front matter.

Regional rule at the level of the statement (manual section 3), 12 and 14 September 2026. A claim is not charted when it names a country outside the region, when it reports a finding from a study the source places in a country outside the region, or when its setting is carried by a regulatory instrument that exists only outside the region (CE marking, the European Union AI Act, the GDPR). Claims about low- and middle-income countries or the BRICS group are claims about groupings to which the region belongs and are charted, labelled pooled_multiregion. A general finding illustrated by a quotation from outside the region is charted.

What is not a barrier (manual section 3, criterion 1), 16 September 2026. Not charted: a study's account of its own limitations; a control already in place; the problem a project was built to solve; heterogeneity described but not asserted as an impediment; adoption status on its own; a paradox framing that names no deficit. An acronym or label the source never expands is not expanded by the coder, and a cell whose deficit the source never states is not a barrier. A sentence naming one deficit that affects several groups is one row. A source can study a policy process empirically without studying the barriers empirically. A primary study's discussion speculating beyond its data is asserted. An expert consensus among a paper's own authors is not data; the evidence rule requires material collected from someone and analysed.

Completeness pass, 16 September 2026. After the first recharting pass, every source was re-read against its charted rows to find what the first-round charting had never captured. Rows added by the pass carry the note "completeness pass"; no restatement rows were added by it. Most additions came from introductory framing and concluding recommendations, which carry citations or no attribution, so the pass moved the corpus toward asserted.

Compliance matrices, 17 September 2026, applied in full 21 September 2026. A table in which a documentary review scores national instruments against a list of principles is charted one row per principle found deficient, never one row per cell, with the countries the cells attribute the gap to recorded in the coder note. Every such row is coded to the deficit the cell asserts (that a national instrument does not cover the principle, a governance and regulatory framework gap) and not to the subject matter of the principle, because the cell does not assert that data are unprotected or models biased in those countries. Applied to Zavaleta-Monestel Table 1: 20 rows for its 20 deficient principles.

Requirement-and-rationale tables, 18 September 2026. In a table pairing a requirement ("the tool should ...") with its rationale, the requirement cell never decides; the rationale cell decides. A row is charted when the rationale states a lack and declined when it states only the benefit of compliance. Applied to Puig-Bosch Table IV.

Remedy-only sentences, 18 September 2026. A recommendation is charted when its own predicate entails a present lack (strengthen, remains, still to be demonstrated, without X Y follows, debe comprobarse) and declined when it names a measure and the goal it serves while predicating nothing of the current state. This is the strict reading, adopted because a rule that let the coder supply an unstated deficit would inflate exactly the quantity the analysis measures.

Adoption status, 18 September 2026. Charted only when the source names what the non-adoption impedes (Dantas R42 stands, Luna R18 and a Chatterjee cell do not).

Borderline rows, 18 September 2026. Where the manual states a default, as the concept exclusion does ("default to excluding"), a borderline row goes out; where it states none, as criterion 1 does, a stated deficit stands.

ai_or_digital is coded from the statement, never from the next sentence or the paper's framing, 18 September 2026.

Contiguous span, 21 September 2026. A quotation may span two contiguous sentences of one paragraph when one claim is carried by both clauses, both quoted whole with no ellipsis across the join (Solari R29). It may never splice text from different structural elements, such as a table cell onto a body sentence.

Regional specificity, rule 4 (manual section 7), 21 and 22 September 2026. A review is the exclusive group when the material it analyses as its evidence is entirely regional; works from outside the region cited only to frame, define or compare do not move it. The test is what the source would lose if the non-regional material were deleted: findings,

and the source is not exclusive; context only, and it is. The separate-arm group requires a separately attributable Latin American result, as an arm, a named country case study or a per-study table row; a source naming a regional country only in a participant table or an exemplar list, with every result pooled, is in the group category.

Verification, 12 and 18 September 2026. The verification mark is empty at charting, Y when the verifier accepts the row, N when the verifier rejects it. Verification is corrective: a row accepted after correction is edited in place and marked Y, so N is reserved for rows rejected outright. No row in the final dataset carries N; the five rows the verifier's decisions removed were dispositioned to the crosswalk or deleted as pass-added rows and are recorded there.

cited_internal is a lower bound, 21 September 2026. The manual resolves a cited claim against the list of included sources. 224 of the 397 cited_external markers are bare numerals ("[29]", "refs 19, 20") that identify nothing outside their own source and were not resolved against the 43 reference lists; the other 173 were scanned for the first author of every included source and none matched. The count of zero is therefore a lower bound and the citation-concentration table (Online Resource 6) is computed on the 173 markers that identify a work.

B2. Controlled vocabulary of sub-themes as it stood when charting closed

The manual fixed 25 sub-themes. The manual's own procedure for admitting a term (propose, adjudicate, sweep every source already coded) was applied fourteen times; each term is named for the deficit the rows state and none supplies content a source does not. Every admission was followed by a sweep of all charted quotations for the new term, and the rows the sweep moved are recorded in the dataset's coder notes.

The 25 terms of the frozen manual. governance/regulatory framework gap; trust/acceptance; equity/access; data quality/availability; connectivity/infrastructure gap; cost/procurement/funding scarcity; digital literacy/workforce capacity gap; workflow integration; data protection/privacy; poor interoperability; workforce shortage/multidisciplinary gap; liability/ethical risk; model validation/generalizability gap; weak collaborative research networks; technological dependency/sovereignty; algorithmic bias/fairness; regulatory heterogeneity/harmonization gap; change management/stakeholder engagement; clinical validation gap; environmental/energy sustainability; explainability/black-box; device/technology approval process; language/cultural adaptation gap; private-sector influence on governance; bureaucratic/administrative barriers.

Terms admitted 16 September 2026.

- *technical performance/accuracy limitation*: the technical limits of the tool itself (Villamarín-Marrugo R17, R18, R19, R21).
- *over-reliance/automation bias/deskilling*: erosion of a clinician's own judgement or skill through reliance on the tool; distinct from technological dependency/sovereignty, which is a system's dependence on foreign technology, and running in the opposite direction to a trust deficit (McLennan R06, Sussman R25).
- *adoption lag/slow uptake* (Zavaleta-Monestel R15).
- *conceptual/terminological imprecision* (Silva R04).
- *analytical/data-processing capacity gap*: the capacity to process and analyse data, distinct from the data's own quality (Chatterjee R02; Bellei R02 recoded to it 18 September).

Terms admitted 17 September 2026.

- *workforce displacement/substitution*: loss of the role, distinct from deskilling, which is erosion of the clinician's skill within the role (Machado p.45 and p.48; Solari p.227; Sánchez-González R16 recoded to it 18 September).
- *evidence transferability gap*: effects demonstrated elsewhere not holding in the setting, a literature-to-setting transfer, distinct from a model's generalizability and from validation in a setting (García Alonso p.63).
- *data access/sharing restriction*: data that exist and are protected but cannot be reached, the opposite deficit to data protection/privacy and not the absence that data quality/availability names (Alencar p.11 and R16).
- *unproven real-world benefit*: benefit not yet demonstrated in use, distinct from whether a model was validated in the setting (Dantas p.367; Gonzales-Aste R25).

- *patient/public AI literacy gap*: lay comprehension of AI information, distinct from workforce digital literacy and from model opacity (Dantas p.368).
- *innovation/R&D capacity gap*: underdeveloped innovation capability (Santa p.2).

Terms admitted 18 September 2026.

- *regional evidence-base under-representation*: the evidence base does not represent the region; it covers deficits scoped to low- and middle-income countries as well as to Latin America, so "regional" in the name is not read as Latin America only (Valli R15, Méndez-Ávila R18, Solari R29, XiquesMolina R07).
- *research funding/agenda misallocation*: the money exists and goes elsewhere, distinct from scarcity and from innovation capacity (McLennan p.10, Yepes R21).
- *health-system fragmentation*: neither the technical deficit that poor interoperability names nor the normative one that regulatory heterogeneity names.

Declined. A generic evidence or evaluation gap term was refused (13 September): a gap in the literature is charted to the sub-theme of what the evidence is missing about, which is what regional evidence-base under-representation later named for the family that grew past five sentences. A usability or interface-design term was declined (18 September) because two of its three candidate sentences named a remedy without stating the deficit.

B3. Locator correction

Page locators were audited mechanically against the layout-preserved text extraction of each PDF (locator_audit.py in the deposit): the extraction is split on the page breaks the extractor writes, each page is scored by how much of the row's quotation it contains, and the printed page number is read from the page furniture. A row is reported only when its charted page scores materially worse than the best page, so a sentence repeated in a table or an abstract raises no false hit. 38 locators across five sources were corrected: Gonzales-Aste, Zavaleta-Monestel and McLennan carried locators one page too low, Solari one page too high, and Silva carried five section labels from a differently numbered chapter. No coded value changed. The direction differed by source because the running page number sits at the foot of some layouts and at the head of others.

B4. Material the extraction could not read

Two figures are image-only in both text extractions and could not be charted: Sánchez-González Figure 3, the per-item breakdown of the concerns paragraph, and Marey Figure 2, the challenges diagram. If either names a barrier its narrative does not, the pass cannot have seen it. Every other table, box and interleaved passage in the 43 sources was read from the layout-preserved extraction.

B5. Verification record

Every row was charted under the manual with the quotation gated mechanically before handover (validate_rechart.py: every quotation must be an exact substring of the source extraction after whitespace and ligature normalisation, or must verify by a sliding-window containment test against the layout-preserved extraction where the two-column layout interleaves a sentence). One reviewer (PPS) verified every row against the full text, editing in place; a second reviewer (JDNS) reviewed the whole dataset; a third (DVJ) audited the count matrices. Verification was corrective rather than adjudicative, so a row accepted after correction was edited in place and no charted row carries a rejection; the corrections themselves are in the deposited dataset. Regional specificity was coded post hoc, one coder per source against the full text under a fixed brief, with the deciding quotation gated against the source extraction in the same way, and verified by PPS; twelve of the 43 codings were flagged borderline and carry both readings in the coder note.

The quotation gate, in numbers. Of the 1,097 charted rows, 1,020 were accepted as an exact substring of the source extraction after whitespace, ligature, hyphenation and elision normalisation. The remaining 77, drawn from 18 sources, are rows whose printed sentence the extraction cannot reproduce as continuous text, because a two-column layout or an interleaved table breaks it; each declares Restored from the PDF in its coder note and was tested instead by sliding-window containment against the layout-preserving extraction of the same PDF, which is the artefact the coder read. Of the 77, 71 verify against that extraction and 6 do not, and they are named here rather than

left in the aggregate: CiecierskiHolmes R02, ElizabethAlburuquequeArana R13, GonzalesAste R06, Luna R04, Soares R02, Sussman R02. This is a queue for a human re-read of the printed page and not a verdict: the benign shape is a sentence the PDF interrupts with a table caption, a footnote block or a hyphenated line break, so the layout stream separates words the printed page prints together. The test ignores word order, which is its blind spot, so a quotation that reverses the printed order passes it; the three rows on which that was found (Sussman R10, R11 and R12) were re-quoted from the printed page. The sources contributing layout-tested rows are ArcosHuaman, Botero, Chatterjee, CiecierskiHolmes, ElizabethAlburuquequeArana, GarciaAlonso, GonzalesAste, Luna, MendezAvila, MorenoFernandez, PuigBosch, Shahid, Soares, Sussman, Valli, VillamarinMarrugo, Werutsky, ZavaletaMonestel.

B6. Reconciliation of the first-round instances

Of the 829 instances charted in the first round, 731 were carried into the recharted dataset (675 as distinct instances and 56 as restatements of one) and 98 were not: 41 fell outside the concept as fixed by the manual (section 3), 37 were attributed to a setting outside the region (regional rule), 16 stated no deficit (criterion 1), and 4 duplicated another instance. Every one of the 98 carries its quotation, locator and the rule applied in the crosswalk file of the deposit. Recharting added 356 distinct instances and 10 restatements the first round had not captured. The final dataset holds 1,097 rows: 1,031 distinct instances and 66 restatements.