Original Article
http://mjiri.iums.ac.ir
Medical Journal of the Islamic Republic of Iran (MJIRI)

Med J Islam Repub Iran. 2026 (11 May);40.46. https://doi.org/10.47176/mjiri.40.46

A Fully Expert Human-Based Retrieval Augmented Generation (FEH-RAG) Framework: A Proof of Concept Study in Labelling Patients with Sjögren Syndrome

Asghar Hajiabbasi[1], Ahmad Reza Jamshidi[2], Shervan Shoaee[3,4,5], Irandokht Shenavar Masooleh[1], Hourvash Ebrahimi[1], Mahsa Mehdipour Dalivand[1], Pooneh Ghavidel Parsa[1]*

Received: 27 Jan 2025 Accepted: 20 Apr 2026 Published: 11 May 2026

Abstract

Background: Accurate application of reference standards (RSs) is essential for correct decision-making in areas governed by such standards. Yet in real-world practice, even fully trained users often apply RSs inconsistently, due to cognitive overload, stress, or other contextual factors, generating misleading evidence. This problem is exemplified by the fact that up to 80% of board-certified hematologists mislabel patients with Sjögren syndrome (SS), a connective tissue disorder (CTD) associated with the greatest risk of lymphoproliferative disorders compared to other CTDs. Embedding RS-based consultations with full objectivity into the decision-making layers of digitalized and non-digitalized settings is critical for improving both decision-making and the quality of resulting evidence. The aim of this proof-of-concept (POC) study is to apply a newly developed framework, the Fully Expert Human-based Retrieval Augmented Generation (FEH-RAG) framework, to provide such a foundation for augmenting SS case labeling in routine daily clinical practice.

Methods: In this POC study, using the nine steps of the FEH-RAG framework, seven expert end-users systematically selected the most widely used SS classification criteria (SSCC) and extracted their elements, including items, definitions, item weights, and inter-item relationships. Extracted items were profiled based on their usage in routine clinical practice. A pathway layout and decision tables were developed accordingly. Following pathway generation, the residual misalignment of the outputs with the SSCC was assessed in a cohort of patients at risk of SS. The experts predefined that the residual misalignment rate of the FEH-RAG outputs with the SSCC must be ≤2% (95% confidence, using the rule of three).

Results: The FEH-RAG framework objectively generated RS-based, transparent, and traceable outputs, including decision tables, an SS classification pathway, and a list of misinterpretations of the SSCC. These misinterpretations involved definitions of dry eye and dry mouth, application of secondary SS criteria, handling SS criteria-specific exclusion rules, and interpretation of serological and objective test results. This POC study achieved its expert-defined maximum misalignment threshold of ≤2% with 95% confidence (0 misalignment in 150 consecutive patients at risk of SS).

Conclusion: This POC study established the needed foundation for improving SS case labeling in daily clinical practice across both digital and non-digital settings. As shown here, publishing FEH-RAG outputs while highlighting potential RS misinterpretations offers a transparent and traceable basis for augmenting decision-making in domains governed by RSs.

Keywords: Error, Sjögren syndrome, Framework, retrieval augmented generation, RAG

Cite this article as: Hajiabbasi A, Jamshidi AR, Shoaee S, Shenavar Masooleh I, Ebrahimi H, Mehdipour Dalivand M, Ghavidel Parsa P. A Fully Expert Human-Based Retrieval Augmented Generation (FEH-RAG) Framework: A Proof of Concept Study in Labelling Patients with Sjögren Syndrome. *Med J Islam Repub Iran.* 2026 (11 May);40:46. https://doi.org/10.47176/mjiri.40.46

Corresponding author: *Dr Pooneh Ghavidel Parsa, grrc@gums.ac.ir*

1. Guilan Rheumatology Research Center, Department of Rheumatology, Razi Hospital, School of Medicine, Guilan University of Medical Sciences, Rasht, Iran
2. Rheumatology Research Center, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran
3. Elderly Health Research Center, Endocrinology and Metabolism Population Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran.
4. Non-communicable Diseases Research Center, Endocrinology and Metabolism Research Institute, Tehran University of Medical Sciences, Tehran, Iran
5. Kerman Oral and Dental Diseases Research Center, Kerman University of Medical Sciences, Kerman, Iran

↑What is "already known" in this topic:
There is a gap between the rules encoded in reference standards and their correct application within real-world settings. This gap is exemplified in Sjögren syndrome (SS), where the absence of diagnostic criteria leads clinicians to rely on SS classification criteria when assigning an SS label. However, clinicians, including expert rheumatologists, frequently misapply these criteria, resulting in a notably high rate of SS patient mislabeling.

→What this article adds:
This POC study shows that the FEH-RAG framework can provide an actionable foundation to augment decision-making in an era governed by reference standards, such as labeling of patients with SS.

[DOI: 10.47176/mjiri.40.46]

Introduction

Accurate application of reference standards (RSs) is fundamental for correct decision-making in settings governed by such standards (1). Yet in real-world practice, even fully trained users often apply RSs inconsistently due to cognitive overload, stress, or other contextual factors (2-7). This not only results in decision errors and inefficient resource utilization but also produces misleading evidence (1, 4, 8-11). This problem is illustrated in the labelling of patients with Sjögren syndrome (SS), a complex connective tissue disease (CTD) with the highest risk of lymphoproliferative disorders (12-14). Clinicians to correctly label a SS case solely rely on the SS classification criteria (SSCC) (12, 13). The SS mislabeling rates, due to the misapplication of the SSCC, even among experienced rheumatologists, are high (15). Up to 80% of board-certified hematologists mislabel the SS cases (14).

Artificial intelligence systems, both data-driven and non-data-driven, are changing the world. They have the potential to reduce human decision errors. However, for their safe daily use in high-stakes domains such as medicine, it is crucial to manage their inherent limitations such as opaque, hallucinated, and biased recommendations (7, 16-20). These challenges cannot be resolved merely by holding their inherently error-prone end users accountable.

Embedding RS-based consultation with full objectivity into the decision-making layers of digitalized and non-digitalized settings is critical for improving decision-making and the quality of the resulting evidence. This can be accomplished via an actionable framework where RS elements, rather than raw evidence, are retrieved objectively, transparently, and traceably, maintaining full fidelity under expert oversight.

To address these gaps, this proof of concept (POC) study applies the fully expert human-based retrieval augmented generation (FEH-RAG) framework, developed by the corresponding author, to the specific challenge of SS labelling.

Methods

In this POC study, to establish an actionable foundation for improving labeling of SS cases by clinicians, digital developers, and regulatory bodies, we applied the FEH-RAG framework.

The FEH-RAG operates through nine iterative steps, both within and between steps (Figure 1). Advancement beyond step 2-7 requires achieving the maximum exact agreement score regarding complete fidelity to the source RSs. This score is calculated as the product of the number of working group (WG) members (n) and the highest Likert scale value (X), ensuring that every member is assigned the top score. Criteria for advancing beyond step 8 are described in the respective section.

Establish a Working Group (Step 1)

A hands-on working group of at least seven to nine members is established (21), comprising RS end users who are concurrently active in domain-specific, educational, and research settings, and who have firsthand experience with real-world workflows and cognitive demands .To maximize the detection of real-world RS misinterpretations, a key output of the framework, it is recommended that all WG members have no more than 15 years of experience in the above settings (22). At least one technical member (TM) is included in the WG. The number of TMs is proportional to the number and complexity of the source RSs. Within this framework, any disagreements arising among TMs while performing their tasks are addressed by the lead TM, as the corresponding or first author.

Selection of the Relevant RSs (Step 2)

To systematically and robustly identify consulted RSs and their validation documents in practice, education, and research, two complementary approaches are employed. First, a field investigation (FI) is conducted. This involves telephone interviews with at least three board-certified experts, independent of the WG, each with a minimum of 15 years of experience and currently actively engaged in domain-specific practice, teaching, and research. Second, RSs relevant to a setting, along with their validation documents, are identified for review. These include standards that are published in domain-specific peer-reviewed journals (e.g., published articles in PubMed, Embase for medicine), incorporated into evidence-based databases (e.g., UpToDate for medicine), or referenced in textbooks.

Review of the Source RSs (Step 3)

Each RS is objectively and iteratively reviewed, first by the TMs and then by the remaining WG members, to extract its elements and generate structured, actionable, and traceable decision tables detailing RS items, their definitions, weights, and inter-item logics. The framework rule of maintaining complete fidelity to the RSs highlights elements prone to misinterpretation. The decision tables should be designed to be directly interpretable and actionable by not only domain-specific users but also other stakeholders such as digital developers and regulatory bodies, with the least possible dependency on domain-specific experts.

Profiling the Items of the RSs (Step 4)

Each RS item is profiled according to its role in decision making within the specific setting in which the FEH-RAG outputs will be used, for example, routine outpatient settings where certain items are not assessed because they are invasive, or limited-resource settings where some items are unavailable.

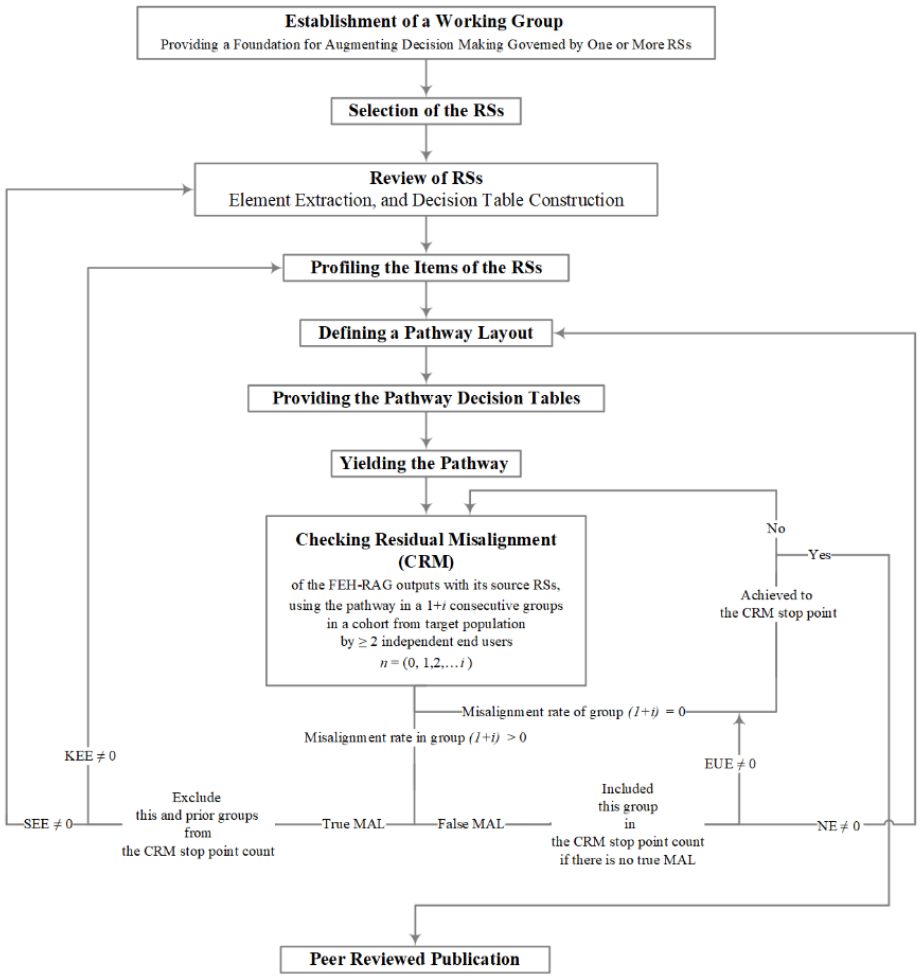
Defining a Pathway Layout (Step 5)

Considering RSs in the context of end users' workflows and cognitive load, the profiled RS items are arranged into hierarchical layers and sublayers to define the pathway layout.

Providing the Pathway Decision Table (Step 6)

For each RS, a top-down decision table is constructed. Items in the first layer and its sublayers are first translated into explicit conditional logic statements, which are then expanded sequentially using items from subsequent layers and sublayers.

[DOI: 10.47176/mjiri.40.46]



TM: Technical member, MAL: Misalignment of the FEH-RAG outputs with RSs, EUE: End users error that is due to misinterpretation of at least one source standard, Stop Point: Defined by the working group using the rule of three, based on real world and targeted FEH-RAG error rates, NE: Navigational error that is due to the navigation of the pathway, KEE: Knowledge engineering error that is due to incorrectly applying correctly extracted elements of the source standards, SEE: Source extraction error that is due to error in extraction of source standards

Figure 1. The Fully Expert Human-Based Retrieval Augmented Generation (FEH – RAG) Framework

Yeilding the Pathway (Step 7)

The lead TM uses the decision tables, the tables of profiled RS items, the pathway layout, and pathway decision tables to develop a preliminary pathway. This pathway is then iteratively refined by the WG.

Checking Residual Misalignment of FEH-RAG Outputs with the Source RSs (Step 8)

To detect residual misalignments (MAL), a cohort drawn from the target population is evaluated by at least two independent end users (2EU) using a paper format of the FEH-RAG pathway and source RSs. A misalignment is recorded when the pathway output does not match the output of at least one of the source RSs.

At this stage, depending on the real-world frequency of decision events and the rate of RS misuse, a prospective, retrospective, or combined cohort may be used. Retrospective cohorts are more appropriate when the decision event is infrequent, enabling efficient use of existing data without requiring prolonged observation, and a prospective cohort is recommended when RS misapplication is common, allowing the outputs to be tested across a wide range of decision scenarios.

Each recorded MAL that is adjudicated and categorized by technical members (TM) is defined as follows:

1. True MAL: The MAL that results from non-fidelity to the source RS in at least one output of FEH-RAG. True MALs may arise from errors in extracting the source RSs (Source Extraction Error: SEE) or from incorrect application of correctly extracted RS elements during steps 4–7 (Knowledge Engineering Error: KEE $_{4\text{-}7}$).

2. False MAL: The MAL that is not due to non-fidelity of at least one FEH-RAG output to the source RSs, but rather to factors related to EUs. These may be due to

[DOI: 10.47176/mjiri.40.46]

misinterpretation of RSs by them (End Users Error: EUE) or suboptimal navigation of the pathway, such as poor layout, excessive length, or visual clutter (Navigational Error: NE).

The cohort is divided into sequential groups, with the size of groups determined dynamically during the study as follows. The size of the initial group is determined by the WG based on key contextual factors, including the prevalence of the decision event (e.g., disease prevalence for diagnostic RSs), the number and complexity of source RSs, and available human and financial resources. The lead of TMs subsequently adjusts the size of the following groups based on the rate and timing of MALs, ensuring efficient use of the cohort by preventing premature or excessive involvement of TMs.

To define the stopping point for this step (Advancement rule) using the rule of three, the WG specifies the maximum acceptable error rate for a decision event after applying the FEH-RAG outputs as a foundation, based on real-world error rate. At this point, the FEH-RAG users will be 95% confident that the rate of true MALs does not exceed the target error rate threshold. For example, suppose that the current error rate for a specific decision is 30%, and the WG decides that, after FEH-RAG implementation, the acceptable error rate should be reduced to 2%. According to the rule of three, 150 consecutive cases without a true MAL are required to provide 95% confidence that the actual true MAL rate does not exceed 2% (5, 23).

Peer-Reviewed Publication of the FEH-RAG-Based Study (Step 9)

To ensure transparency of both the FEH-RAG outputs and the systems built upon them, publication in a peer-reviewed journal is mandated as the final, non-optional step of the framework. This enables stakeholders to critically appraise the FEH-RAG outputs.

To establish a foundation for augmenting labeling of SS cases, we implemented the FEH-RAG as follows:

IRAPSS WG Establishment (Step 1)

We convened a multidisciplinary WG, comprising members who were actively engaged in teaching, research, and clinical practice at Iran medical universities including three rheumatologists with more than 25 years expertise as rheumatologist, three rheumatologists with more than 2 years expertise as rheumatologist, one dentist with more than 25 years expertise as dentist and 10 years as epidemiologist, a general practitioner with more than 25 years of experience including over six years of service as an executive and clinical member of the Iran Primary Sjögren Syndrome Registry (IRAPSS) (24).

Selection Relevant RSs for Labelling SS Patients (Step 2)

The only available guidance for correctly classifying a patient with SS is provided by the SSCCs. To determine which SSCCs are most widely used across clinical practice, teaching, and research, and to identify their original validation articles, in which each SSCC was first introduced and its validation methodology described, we reviewed the following resources:

a. Field investigation (Fi): We conducted telephone interviews with five board-certified rheumatologists, each with more than 15 years of clinical, teaching, and research experience. Among them, three were also members of the IRAPSS registry.

b. Rheumatology textbooks: The diagnostic sections of widely used rheumatology textbooks in medical education, including Kelley's Textbook of Rheumatology and Hochberg's Rheumatology, were reviewed.

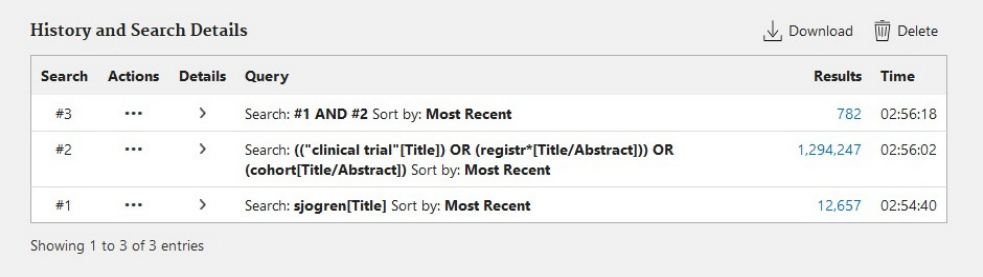
c. Evidence-based medicine resource: The UpToDate platform was assessed to identify the SSCC referenced in it.

d. PubMed database: A comprehensive literature review was performed to identify the SSCCs applied in SS registries and clinical trials. Screening was conducted based on titles and/or abstracts, followed by full-text review when necessary (Figure 2).

The three most commonly used SSCCs were the American European Consensus Group (AECG) SS Classification Criteria (25) , the American College of Rheumatology/European League Against Rheumatism (ACR/ EULAR) Primary Sjögren Syndrome Classification Criteria (26), and the American College of Rheumatology (ACR) SS Classification Criteria (27). These criteria are supported by the highest available levels of evidence, derived primarily from large international cohort studies. To maintain full fidelity to the SSCC, the SSCC's original validation articles were used as references for resolving any disagreements within the WG.

Reviewing The SSCC (Step 3)

For the two categories of the SSCC, primary SS (pSS) and secondary SS (sSS), five extraction tables (Tables 1-5) were developed to represent all SSCC items, definitions, item weights, and inter-item logical rules.


History and Search Details

Download Delete

Search	Actions	Details	Query	Results	Time
#3	•••	›	Search: **#1 AND #2** Sort by: **Most Recent**	782	02:56:18
#2	•••	›	Search: **(("clinical trial"[Title]) OR (registr*[Title/Abstract])) OR (cohort[Title/Abstract])** Sort by: **Most Recent**	1,294,247	02:56:02
#1	•••	›	Search: **sjogren[Title]** Sort by: **Most Recent**	12,657	02:54:40

Showing 1 to 3 of 3 entries

Figure 2. PubMed Search Strategy for Selecting Sjögren Syndrome Classification Criteria

[DOI: 10.47176/mjiri.40.46]

Table 1. Decision table of 2002 American European Consensus Group criteria (AECG) for Sjögren syndrome, in respect of primary Sjögren syndrome (pSS) (25)

	Term of use	Items that should be excluded	Ophthalmic symptoms	Oral symptoms	Anti Ro Ab	Anti La Ab	Schirmer test without analgesic	Rose Bengal or other ocular dye score	Unstimulated whole salivary flow rate test	Parotid sialography	Salivary scintigraphy	Minor salivary gland biopsy
Definitions	Classifying a connective tissue disease case as primary Sjögren syndrome	Hepatitis C infection AIDS[1] Sarcoidosis GVHD[2] Past head and neck radiation treatment Preexisting lymphoma Anticholinergic drugs[3]	Positive answer to at least one of the three items: 1. Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye) 2. Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes) 3. Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less and not only ophthalmic drops)	Positive answer to at least one of the three items: 1. Having a daily feeling of dry mouth for MORE THAN 3 MONTHS (not just a feeling of dry mouth) 2. FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food) 3. RECURRENTLY or PERSISTENTLY swollen salivary glands in ADULT	+ means: Above normal limit	+ means: Above normal limit	+ means: ≤ 5 mm in 5 minutes	+ means: ≥ 4 according to van Bijesterveld's score	+ means: ≤ 1.5 cc per 15 minutes or ≤ 0.5 cc per 5 minutes	+ means: The presence of diffuse sialectasias (punctate, cavitary or destructive pattern), without evidence of obstruction in the major ducts.	+ means: Delay uptake, reduced concentration and/or delay excretion of tracer	+ means: Obtained through normal-apperaing mucousa. Focal lymphocytic sialoadenitis, evaluated by an expert histologist, with a focus score ≥1, defined as a number of lymphocytic foci (which are adjusted to normal-appearing mucous acini and contain more than 50 lymphocytes) per 4 mm^2 of glandular tissue
Weight Rules			If negative, = 0. If positive + but Ro[4], La[5] and MSGB[6] are - , = 0. If positive and Ro, La and/or MSGB are/is +, = 1.	If negative, = 0. If positive + but Ro, La and MSGB are - , = 0. If positive and Ro, La and/or MSGB are/is +, = 1.	1 if Ro and/or La is +. 0 if both are -		If all of them is -, = 0. If at least one of them is + and Ro, La and MSGB are - , = 0. If at least one of them is + and Ro, La or MSGB are/is +, = 1.		If all of them is -, = 0 If at least one of them is + and Ro, La and MSGB are -, = 0. If at least one of them is + and Ro, La or MSGB are/is +, = 1			If - = 0 If + = 1
							If sum of weight ≥ 4 → AECG is met for pSS. In pateints without ophthalmic and oral symptoms: If sum of weight ≥ 3 → AECG is met for pSS					

1.AIDS: acquired immunodeficiency syndrome,2.Graft versus host disease, 3. Use of anticholinergic drugs for a time shorter than 4 times the half-life of life of drug,4. Anti Ro antibody, 5. Anti La, 6. Minor salivary gland biopsy

[DOI: 10.47176/mjiri.40.46]

Table 2. Decision table of 2002 American European Consensus Group criteria (AECG) for Sjögren syndrome, in respect of secondary Sjögren syndrome (sSS) (25)

	Term of use	Items that should be excluded	Schirmer test without analgesic	Rose Bengal or other ocular dye score	Unstimulated whole salivary flow rate test	Parotid sialography	Salivary scintigraphy	Minor salivary gland biopsy
Definitions	[An individual who has: oral symptoms (Positive response to at least one of them, including: 1. Having a DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just a feeling of dry mouth) 2 .FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food), 3. RECURRENTLY or PERSISTENTLY swollen salivary glands in an ADULT, and/or ophthalmic symptoms (Positive response to at least one of them, including: 1 .Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), 2 . Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), 3 .Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less and not only ophthalmic drops),]	Hepatitis C infection AIDS[1] Sarcoidosis GVHD[2] Past head and neck radiation treatment Preexisting lymphoma Anticholinergic drugs[3]	+ means: ≤ 5 mm in 5 minutes	+ means: ≥ 4 according to van Bijesterveld's score	+ means: ≤ 1.5 cc per 15 minutes	+ means: The presence of diffuse sialectasias (punctate, cavitary or destructive pattern), without evidence of obstruction in the major ducts.	+ means: Delay uptake, reduced concentration and/or delay excretion of tracer	+ means: Obtained through normal-appearing mucousa. Focal lymphocytic sialoadenitis, evaluated by an expert histologist, with a focus score ≥1, defined as a number of lymphocytic foci (which are adjusted to normal-appearing mucous acini and contain more than 50 lymphocytes) per 4 mm^2 of glandular tissue
Weight Rules	and Have CORRECTLY MET the classification criteria of another well-defined connective tissue disease, such as systemic lupus erythematosus or rheumatoid arthritis etc.		If all of them is -, = 0. If at least one of them is +, = 1		If all of them is -, = 0. If at least one of them is +, = 1			If - = 0 If + = 1
			If sum of weights ≥ 2→ AECG is met for sSS.					

1. AIDS: acquired immunodeficiency syndrome,2.Graft versus host disease, 3.Use of anticholinergic drugs since a time shorter than 4 fold the half of life of drug.

[DOI: 10.47176/mjiri.40.46]

Table 3. Decision table of 2016 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) for primary Sjögren syndrome (pSS) criteria (26)

	Term of use	Items that should be excluded	Anti Ro Ab	Schirmer test without analgesic	Ocular staining score	Van Bijsterveld score	Unstimulated whole salivary flow rate test	Minor salivary gland biopsy
Definitions	[An individual positive response to at least 1 of the following questions 1 .Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye) 2 .Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes) 3 .Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less and not only ophthalmic drops) 4. Having a DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just a feeling of dry mouth) 5 .FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food) OR A patient with other SS suggestive symptoms or signs based on Eular Sjögren syndrome disease activity index] AND [The individual HAS NOT CORRECTLY MET the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.]	Active hepatitis C infection (with confirmation by polymerase chain reaction) AIDS[1] Sarcoidosis GVHD[2] History of head and neck radiation treatment Amyloidosis IgG4 related disease Anticholinergic drugs[3]	+ means: Above normal limit	+ means: ≤ 5 mm in 5 minutes	+ means: ≥ 5	+ means: ≥ 4	+ means: ≤ 0.5 cc per 5 minutes	+ means: Obtained through normal-apperaing mucousa. Focal lymphocytic sialoadenitis, evaluated by an expert histologist, with a focus score ≥1, defined as a number of lymphocytic foci (which are adjusted to normal-appearing mucous acini and contain more than 50 lymphocytes) per 4 mm^2 of glandular tissue
Weight Rules			If - = 0 If + = 3	If all of them is -, = 0. If at least one of them is +, = 1			If - = 0 If + = 1	If - = 0 If + =3
			If sum of weights ≥ 4 → ACR/EULAR is met for pSS.					

1.AIDS: acquired immunodeficiency syndrome,2.Graft versus host disease, 3.Use of anticholinergic drugs in previous 24 hours.

[DOI: 10.47176/mjiri.40.46]

Table 4. Decision table of 2012 American College of Rheumatology criteria (ACR) for Sjögren syndrome, in respect of primary Sjögren syndrome (pSS) (27)

	Term of use	Items that should be excluded	Anti Ro Ab	Anti La Ab	Antinuclear Ab	Rheumatoid factor	Keratoconjectivitis SICCA score	Minor salivary gland biopsy
Definitions	An individual with SS suggestive symptoms or signs who HAS NOT CORRECTLY MET the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.	Hepatitis C infection AIDS[1] Sarcoidosis GVHD[2] History of head and neck radiation treatment Amyloidosis IgG4 related disease Anticholinergic drugs[3]	+ means: Above normal limit	+ means: Above normal limit	+ means: Above normal limit by IMMUNOFLUORESCENCE method	+ means: Above normal limit	+ means: ≥ 3, assuming that individual is not currently using daily eye drops for glaucoma, and has not had corneal surgery or cosmetic eyelid surgery in the last 5 years)	+ means: Obtained through normal apperaing mucousa. Focal lymphocytic sialoadenitis, evaluated by an expert histologist, with a focus score ≥1, defined as a number of lymphocytic foci (which are adjusted to normal-appearing mucous acini and contain more than 50 lymphocytes) per 4 mm^2 of glandular tissue
Weight Rules			0 if Ro – and La – and, one of ANA or RF -. 1 if (Ro[4] or La[5]) or, (ANA[6] and RF[7]) was +				0 if -. 1 if +	0 if -. 1 if +
			If sum of weights ≥ 2 → ACR is met for pSS.					

1.AIDS: acquired immunodeficiency disease syndrome,2.Graft versus host disease, 3.Use of anticholinergic drugs in previous 24 hours,4. Anti Ro antibody, 5. Anti La, 6. Antinuclear antibody with immunofluorescence method, 7. Rheumatoid factor

Table 5. Decision table of 2012 American College of Rheumatology criteria (ACR) for secondary Sjögren syndrome (sSS) (27)

	Term of use	SS items that should be excluded	Anti Ro Ab	Anti La Ab	Antinuclear Ab	Rheumatoid factor	Keratoconjectivitis SICCA score	Minor salivary gland biopsy
Definitions	An individual with suggestive SS symptoms or signs who HAS CORRECTLY MET the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.	Hepatitis C infection AIDS[1] Sarcoidosis GVHD[2] Head and neck radiation treatment Amyloidosis IgG4 related disease Anticholinergic drugs[3]	+ means: Above normal limit	+ means: Above normal limit	+ means: Above normal limit by IMMUNOFLUORESCENCE method	+ means: Above normal limit	+ means: ≥ 3, assuming that the individual is not currently using daily eye drops for glaucoma, and has not had corneal surgery or cosmetic eyelid surgery in the last 5 years)	+ means: Obtained through normal-apperaing mucousa. Focal lymphocytic sialoadenitis, evaluated by an expert histologist, with a focus score ≥1, defined as a number of lymphocytic foci (which are adjusted to normal-appearing mucous acini and contain more than 50 lymphocytes) per 4 mm^2 of glandular tissue
Weight Rules			0 if Ro – and La – and, one of ANA or RF -. 1 if (Ro[4] or La[5]) or, (ANA[6] and RF[7]) was +				0 if -. 1 if +	0 if -. 1 if +
			If sum of weights ≥ 2 → ACR is met for sSS.					

1.AIDS: acquired immunodeficiency syndrome,2.Graft versus host disease, 3.Use of anticholinergic drugs in previous 24 hours,4. Anti Ro antibody, 5. Anti La, 6. Antinuclear antibody with immunofluorescence method, 7. Rheumatoid factor

[DOI: 10.47176/mjiri.40.46]

Profiling of the SSCC Items (Step 4)

The items of the SSCC were profiled according to their utility in the clinicians' routine clinical practice (Table 6):

- Type: Clinical versus paraclinical
- Accessibility: Available, hardly available, Not available
- Invasiveness: Not invasive, invasive, very invasive
- Reliability: The most, the least
- Use in clinical routines: Yes, No

Laying out the IRAPSS Pathway (Step 5)

First, the minimum data set (MDS) for the IRAPSS pathway was selected (bolded items in Table 6). These items, routinely assessed in daily clinical practice, are widely available and highly reliable and include anti-Ro and anti-La antibodies, antinuclear antibody assessed by immunofluorescence (FANA), rheumatoid factor, Schirmer's test without analgesic (Schirmer), unstimulated whole salivary flow rate (USWSFR), and minor salivary gland biopsy (MSGB), together with the inclusion and exclusion rules defined by the SSCCs.

Then, using the source extraction tables (Tables 1–5) and the MDS (Table 6), the hierarchical architecture of the IRAPSS pathway was established (Table 7), incorporating layers, sublayers, and the structured positioning of each item. In designing this pathway, it was assumed that patients either had no prior rheumatologic diagnosis or that any previously assigned rheumatologic diagnoses were incorrect. This approach allows identification of SS patients who may have been previously mislabeled, ensuring accurate labeling independent of prior diagnostic errors.

Accurate labeling also requires careful consideration of potential false-positive and false-negative paraclinical results. To address this, '+' and '−' were defined as true positive and true negative, respectively. For example, a negative anti-Ro result may conceal low-titer anti-Ro52 positivity, while a negative Schirmer's test could result from prior analgesic use. Likewise, unstandardized methodologies may produce misleading USWSFR or false-negative MSGB results.

To facilitate access to the standard method of USWSFR, Schirmer and MSGB, and Eular Sjögren syndrome disease activity index (ESSDAI), four QR codes linking to them were provided alongside the IRAPSS pathway.

IRAPSS Pathway Decision Tables (Step 6)

The IRAPSS pathway decision tables were developed based on clinical workflows (Tables 8–13). During their construction, it was assumed that a patient met inclusion rules and have no any exclusion criteria. Additional assumptions were as follows:

a. pSS section: Patients were assumed to have no well-defined concurrent CTD.

b. sSS section: Patients were assumed to have a well-defined concurrent CTD.

Yeilding The IRAPSS Pathway (Step 7)

The TM developed a preliminary pathway using the SSCC decision tables, the tables of profiled items, the IRAPSS pathway layout, and pathway decision tables. This pathway was then iteratively refined by the WG.

Checking Residual Misalignment to the SSCC (Step 8)

Using a cohort of patients at risk of SS drawn from the IRAPSS registry clinic, two independent clinicians evaluated whether the labeling of patients, using the paper-based SSCC and pathway, was equal. The detected misalignments (MALs) by them were adjudicated and categorized by a TM:

a. All MALs were false.

b. All false MALs were due to NE secondary in the pathway on multiple pages.

To determine the stopping point for this step, the WG considered a maximum acceptable true MAL rate of 2% with 95% confidence. To meet this threshold, it was calculated that at least 150 consecutive patients without true MALs were required.

The stopping criterion was satisfied after 150 consecutive patients with no true MALs. The step was continued as part of quality assurance for patient enrollment in the IRAPSS registry. Among the subsequent 60 consecutive patients, no SEE, KEE, or NE events were observed.

Peer-reviewed Publication Study (Step 9)

This study was published in a journal that is indexed in PubMed.

Results

To provide an actionable foundation for augmenting labeling of cases with SS, we used the FEH-RAG framework. This framework provided not only the SSCC decision tables but also the profiling, layout, and decision tables of the IRAPSS pathway. Based on FEH-RAG, progression through steps 2–7 required attaining the maximum exact-agreement score about complete fidelity to the SSCC. Achieving this enabled the identification of a list of SSCC misinterpretations, which included (25-27):

1. Definitions of the SS-related dry eye, and dry mouth:

A case is considered to have SS related dry eye only if dry eye exists daily, for more than three months AND, is persistent and troublesome OR results in recurrent sensation of sand or gravel in eyes OR induces the need to use at least 4 times per day tear substitute to recover (Ask about the times needed for dry eye recovery NOT the frequency of tear substitute by the patient). In addition, not all tear substitutes are ophthalmic drops. Ophthalmic lubricants should be counted.

A case has SS-related dry mouth, only if dry mouth exists daily and for more than three months, OR results in the need to drink liquids to help in swallowing dry food and it occurs frequently.

2. Definition for the SS-related salivary gland swelling:

Recurrent or persistent salivary gland swelling, which is not due to any other causes such as stone, infection, IgG4-related disease, sarcoidosis, and so on, is considered SS-related only if it has happened during the adulthood period.

[DOI: 10.47176/mjiri.40.46]

Table 6. Profile of Sjögren syndrome classification criteria items (25-27) based on their use in routine clinical practice (Bolded items were selected for using in the IRAPSS pathway)

		Item characteristics				
		Type	Availability	Invasiveness	Reliability	Requested in clinicians' clinical routines
Items of selected Sjögren syndrome classification criteria	**SS suggestive symptoms and signs (Including Eular Sjögren syndrome disease activity index)**	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Oral and ophthalmic symptoms	Clinical	Available	Not invasive	The most	Yes
	Hepatitis C infection	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Active Hepatitis C infection was confirmed by polymerase chain reaction	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Acquired immunodeficiency syndrome	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Amyloidosis	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Sarcoidosis	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Graft Versus host disease	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	IgG 4 related disease	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	History of head and neck radiation treatment	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Preexisting lymphoma	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes
	Using anticholinergic drugs	Clinical	Available	Not invasive	The most	Yes
	Anti Ro Ab	Para-clinical	Available	Invasive	The most	Yes
	Anti La Ab	Para-clinical	Available	Invasive	The most	Yes
	Rheumatoid factor	Para-clinical	Available	Invasive	The most	Yes
	Antinuclear antibody (by immunofluorescence method)	Para-clinical	Available	Invasive	The most	Yes
	Unstimulated whole salivary flow rate	Para-clinical	Available	Non invasive	The most	Yes
	Parotid sialography	Para-clinical	Hardly available	Very invasive	The least	No
	Salivary scintigraphy	Para-clinical	Hardly available	Very invasive	The least	No
	Schirmer test without analgesic	Para-clinical	Available	Invasive	The most	Yes
	Ocular Staining Score (SICCA or other dye score)	Para-clinical	Not Available	Invasive	The most	No
	Minor salivary gland biopsy	Para-clinical	Available	Very Invasive	The most	Yes
	Do your patients who have fulfilled at least one of the primary Sjögren syndrome criteria, have correctly met the classification criteria of another well-defined connective tissue disease, such as systemic lupus erythematosus or rheumatoid arthritis etc?	Clinical and or para-clinical	Available	Not to very invasive	The most	Yes

[DOI: 10.47176/mjiri.40.46]

Table 7. Layout of the IRAPSS pathway (layers, sublayers and, item list (25-27) and location)

Outlines	Layer name	Rank	Sub-layer name	Sub layers ranks
Check the primary Sjögren criteria (pSS)	Individual with suspected or confirmed connective tissue disorder	0	---	----
	Terms of use and excluding rules	1	Oral [1] and Ophthalmic[2] symptoms and using anticholinergic drug/s	1
			Other SS suggestive symptoms or signs	2
		2	Conditions need to be excluded: List [3] and using anticholinergic drug/s	----
	Paraclinics required to determine fulfillment of pSS classification criteria	3	Anti Ro Ab, Anti La Ab, Antinuclear Ab by immunofluorescence method, rheumatoid factor	1
			Schirmer test without analgesic and Unstimulated whole salivary flow rate and using anticholinergic drug/s	2
			Minor salivary gland biopsy	3
			Type of pSS criteria that was met	4
	Assessment of whether the case correctly meets classification criteria for another well-defined connective tissue disease (e.g., systemic lupus erythematosus or rheumatoid arthritis)	4	---	---
Check the secondary Sjögren criteria (sSS)	Term of use	5	Correctly met the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.	1
			Checking Oral and/or Ophthalmic symptoms AND using anticholinergic drug/s AND whether ACR was met	2
		6	Conditions need to be excluded: List [3]	---
	Paraclinics required to determine fulfillment of sSS classification criteria	7	Schirmer test without analgesic and Unstimulated whole salivary flow rate and using anticholinergic drug/s	1
			Minor salivary gland biopsy	2
			Type of sSS criteria that was met	3

1. Oral symptoms: Positive response to at least one of them, including: a. Having a DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just a feeling of dry mouth) b .FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food), c. RECURRENTLY or PERSISTENTLY swollen salivary glands in an adult, 2. Ophthalmic symptoms: Positive response to at least one of them including: a .Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), b .Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), c. Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less), 3. a. Active hepatitis C (not only infection of hepatitis C): positive polymerase chain reaction, b. Infection of hepatitis c. Acquired immunodeficiency syndrome, d. Amyloidosis, e. Sarcoidosis, f. Graft versus host disease, g. IgG 4 related disease, h. History of head and neck radiation treatment, i. Preexisting lymphoma.

[DOI: 10.47176/mjiri.40.46]

Table 8. Decision table of primary Sjögren syndrome (pSS) section of IRAPSS pathway; item, definitions, weight of items and, inter item logics (continued):

	Term of use	SS items that should be excluded	Ophthalmic symptoms[4]	Oral Symptoms[8]	Ro	La	RF[9]	ANA[10]	Schirmer [11]	USWSFR[12]	MSGB
2002 American European Consensus Group criteria for primary Sjögren syndrome (AECG) (25)											
Weight Rules	Classifying a connective tissue disease case as primary Sjögren syndrome	Hepatitis C infection AIDS[1] Sarcoidosis GVHD[2] History of head and neck radiation treatment Preexisting lymphoma Anticholinergic drugs[3]	If negative, = 0. If positive + but Ro[5], La [6] and MSGB[7] are - , = 0. If positive and Ro, La and/or MSGB are/is +, = 1	If negative, = 0. If positive + but Ro, La and MSGB are - , = 0. If positive and Ro, La and/or MSGB are/is +, = 1	1 if Ro and/or La is +. 0 if both are -		0	0	If Schirmer is -, = 0. If Schirmer is + and Ro, La and MSGB are -, = 0 If Schirmer is + and Ro, La or MSGB are/is +, = 1	If USWSFR is -, = 0. If USWSFR is + and Ro, La and MSGB are -, = 0 If USWSFR is + and Ro, La or MSGB are/is +, = 1	If - = 0 If + = 1
			If sum of weight ≥ 4 → AECG is met for pSS. In pateints without ophthalmic and oral symptoms: If sum of weight ≥ 3 → AECG is met for pSS								
2016 American College of Rheumatology/European League Against Rheumatism criteria (ACR/EULAR) (26)											
Weight Rules	[An individual positive response to at least 1 of the following questions 1 .Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), 2 . Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), 3 .Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less), 4. Having a daily feeling of dry mouth for MORE THAN 3 MONTHS (not just a feeling of dry mouth), 5 .FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food)] OR [A patient with other SS suggestive symptoms or signs based on Eular Sjögren] AND [The individual HAS NOT CORRECTLY MET the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.]	Active hepatitis C[13] AIDS Sarcoidosis GVHD History of head and neck radiation treatment Amyloidosis IgG4 related disease Anticholinergic drugs	0	0	0 if Ro-. 3 if Ro+	0	0	0	If Schirmer is -, = 0. If Schirmer is +, = 1	If USWSFR is -, = 0. If USWSFR is +, = 1	If - = 0 If + = 3
			If sum of weights ≥ 4 → ACR/EULAR is met for pSS.								

1.AIDS: Acquired immunodeficiency syndrome,2.Graft versus host disease, 3.Use of anticholinergic drugs for a time shorter than 4 fold the half-life of the drug, 4. Ophthalmic symptoms definition- Positive response to at least one of them including: a. Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), b. Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), c. Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less), 5. Anti Ro antibody, 6. Anti La, 7. Minor salivary gland biopsy, 8 Oral symptoms definition – Positive response to at least one of them including: a. Having DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just feeling of dry mouth) b. FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food), c. RECURRENTLY or PERSISTENTLY swollen salivary glands in ADULT, 9. Rheumatoid factor, 10.Antinuclear antibody with immunofluorescence method, 11. Schirmer test without analgesic, 12. USWSFR: Un stimulated whole salivary flow rate, 13. Confirmed by polymerase chain reaction.

[DOI: 10.47176/mjiri.40.46]

Table 8. Decision table of primary Sjögren syndrome (pSS) section of IRAPSS pathway; item, definitions, weight of items and, inter item logics

	2012 American College of Rheumatology criteria (ACR) (27)										
	Term of use	SS items that should be excluded	Ophthalmic symptoms	Oral symptoms	Ro	La	RF	ANA	Schirmer	USWSFR	MSGB
Weight Rules	An individual with SS suggestive symptoms or signs who HAS NOT CORRECTLY MET the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.	Hepatitis C infection AIDS Sarcoidosis GVHD History of head and neck radiation treatment Amyloidosis IgG4 related disease Anticholinergic drugs	0	0	If Ro- and La- and, one of ANA or RF - = 0 If (Ro+ and or La+) or (ANA+ and RF +) =1				0	0	If - = 0 If + = 1
			If sum of weights ≥ 2 → ACR is met for pSS.								

1.AIDS:Acquired immunodeficiency syndrome,2.Graft versus host disease, 3.Use of anticholinergic drugs since a time shorter than 4 times the half-life of drug, 4. Ophthalmic symptoms definition- Positive response to at least one of them including: a. Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), b. Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), c. Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less), 5. Oral symptoms definition – Positive response to at least one of them including: a. Having DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just feeling of dry mouth) b. FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food), c. RECURRENTLY or PERSISTENTLY swollen salivary glands in ADULT, 6. Anti Ro antibody, 7. Anti La, 8. Rheumatoid factor, 9. Antinuclear antibody with immunofluorescence method, 10. Schirmer test without analgesic, 11. USWSFR: Un stimulated whole salivary flow rate, 12.Minor salivary gland biopsy,

[DOI: 10.47176/mjiri.40.46]

Table 9. Decision table of secondary Sjögren syndrome (sSS) section of IRAPSS pathway; item, definitions, weight of items and, inter item logics

	Term of use	SS items that should be excluded	Oral symptoms	Ophthalmic symptoms	Ro[6]	La[7]	RF[8]	ANA[9]	Schirmer [10]	USWSFR[11]	MSGB[12]
2002 American European Consensus Group criteria for Sjögren syndrome (AECG) (25)											
Weight	A case with oral[1] and or ophthalmic[2] symptoms that CORRECTLY MEETs the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.	Hepatitis C infection AIDS[3] Sarcoidosis GVHD[4] History of head and neck radiation treatment Preexisting lymphoma Anticholinergic drugs[5]	0	0	0	0	0	0	If Schirmer is -, = 0. If Schirmer is +, = 1	If USWSFR is -, = 0. If USWSFR is +, = 1	If - = 0 If + = 1
Rules			If sum of weights ≥ 2 → AECG is met for sSS.								
2012 American College of Rheumatology criteria (ACR) (27)											
Weight	An individual with suggestive SS symptoms or signs who HAS CORRECTLY MET the classification criteria of another well-defined connective tissue disease such as systemic lupus erythematosus or rheumatoid arthritis etc.	Hepatitis C infection AIDS Sarcoidosis GVHD Head and neck radiation Preexisting lymphoma Amyloidosis IgG4 related disease Anticholinergic drugs	0	0	If Ro- and La- and, one of ANA or RF - = 0 If (Ro+ and or La+) or (ANA+ and RF +) =1				0	0	If - = 0 If + = 1
Rules			If sum of weights ≥ 2 → ACR is met for sSS.								

1. Oral symptoms definition – Positive response to at least one of them including: a. Having a DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just a feeling of dry mouth) b. FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food), c. RECURRENTLY or PERSISTENTLY swollen salivary glands in ADULT, 2. Ophthalmic symptoms definition- Positive response to at least one of them including: a. Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), b. Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), c. Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less), 3.AIDS:Acquired immunodeficiency syndrome,4.Graft versus host disease, 5.Use of anticholinergic drugs since a time shorter than 4 fold the half of life of drug, 6. Anti Ro antibody, 7. Anti La, 8. Rheumatoid factor, 9. Antinuclear antibody with immunofluorescence method, 10. Schirmer test without analgesic, 11. USWSFR: Un stimulated whole salivary flow rate, 12. Minor salivary gland biopsy.

[DOI: 10.47176/mjiri.40.46]

Table 10. Scenarios based decision table of American Eular Consensus Group criteria (AECG) for primary Sjögren syndrome (pSS); combination of oral symptoms, ophthalmic symptoms, Schirmer test without analgesic, unstimulated whole salivary flow rate and minor salivary gland biopsy and their potential overlap with its secondary Sjögren syndrome (sSS) criteria (continued) (25)

Possible scenarios	Oral[1]	Ophthalmic[2]	Serology[3]	Schirmer[4]	USWSFR[5]	MSGB[6]	Met rule: 4/6, 3 objective/ No pSS	Potential overlap with AECG - defined sSS (assuming existence of another CTD) (sSS, m.sSS: may be sSS if MSGB+, or No sSS)
1			+	+	+	No Need	4/6, 3 objectives	sSS
2			+	+	-	No Need	4/6, 2 objectives	m.sSS
3			+	-	+	No Need	4/6, 2 objectives	m.sSS
4			+	-	-	Mandatory +	4/6, 2 objectives	No sSS
	+	+				Mandatory -	No pSS	No sSS
5			-	+	+	Mandatory +	4/6, 3 objectives	sSS
						Mandatory -	No pSS	sSS
6			-	+	-	Mandatory +	4/6, 2 objectives	sSS
						Mandatory -	No pSS	No sSS
7			-	-	+	Mandatory +	4/6, 2 objectives	sSS
						Mandatory -	No pSS	No sSS
8			-	-	-	No Need	No pSS	No sSS
9			+	+	+	No Need	4/6, 3 objectives	sSS
10			+	+	-	Mandatory +	4/6, 3 objectives	sSS
						Mandatory -	No pSS	No sSS
11			+	-	+	Mandatory +	4/6, 3 objectives	sSS
						Mandatory -	No pSS	No sSS
12	+	-	+	-	-	No Need	No pSS	No sSS
13			-	+	+	Mandatory +	4/6, 3 objectives	sSS
						Mandatory -	No pSS	sSS
14			-	+	-	No Need	No pSS	m.sSS
15			-	-	+	No Need	No pSS	m.sSS
16			-	-	-	No Need	No pSS	No sSS
17			+	+	+	No Need	4/6, 3 objectives	sSS
18			+	+	-	Mandatory +	4/6, 3 objectives	sSS
						Mandatory -	No pSS	No sSS
19			+	-	+	Mandatory +	4/6, 3 objectives	sSS
						Mandatory -	No pSS	No sSS
20	-	+	+	-	-	No Need	No pSS	No sSS
21			-	+	+	Mandatory +	4/6, 3 objectives	sSS
						Mandatory -	No pSS	sSS
22			-	+	-	No Need	No pSS	m.sSS
23			-	-	+	No Need	No pSS	m.sSS
24			-	-	-	No Need	No pSS	No sSS

[DOI: 10.47176/mjiri.40.46]

Table 10. Scenarios-based decision table of American Eular Consensus Group criteria (AECG) for primary Sjögren syndrome (pSS); combination of oral symptoms, ophthalmic symptoms, Schirmer test without analgesic, unstimulated whole salivary flow rate and minor salivary gland biopsy and their potential overlap with its secondary Sjögren syndrome (sSS) criteria (25)

Possible scenarios	Oral[1]	Ophthalmic[2]	Serology[3]	Schirmer[4]	USWSFR[5]	MSGB[6]	Met rule: 4/6, 3 objective/ No pSS	Potential overlap with AECG - defined sSS (assuming existence of another CTD) (sSS, m.sSS: may be sSS if MSGB+, or No sSS)
25			+	+	+	No Need	3 Objectives	No sSS
26			+	+	-	Mandatory +	3 Objectives	No sSS
						Mandatory -	No pSS	No sSS
27			+	-	+	Mandatory +	3 Objectives	No sSS
	-	-				Mandatory -	No pSS	No sSS
28			+	-	-	No Need	No pSS	No sSS
29			-	+	+	Mandatory +	3 Objectives	No sSS
						Mandatory -	No pSS	No sSS
30			-	+	-	No Need	No pSS	No sSS
31			-	-	+	No Need	No pSS	No sSS
32			-	-	-	No Need	No pSS	No sSS

1.Oral symptoms definition – Positive response to at least one of them including: a. Having a DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just feeling of dry mouth) b. FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food), c. RECURRENTLY or PERSISTENTLY swollen salivary glands in ADULT, 2. Ophthalmic symptoms definition- Positive response to at least one of them including: a. Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), b. Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), c. Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less), 3. Serology: Positive anti Ro Ab and/or anti La Ab, 4. Schirmer: Schirmer test without analgesic, 5. USWSFR: Unstimulated whole salivary flow rate, 6. MSGB: Minor salivary gland biopsy.

Table 11. Scenarios based decision table of American Eular Consensus Group criteria (AECG) for secondary Sjögren syndrome (sSS); combination of oral symptoms, ophthalmic symptoms, Schirmer test without analgesic, unstimulated whole salivary flow rate and minor salivary gland biopsy (25)

Possible scenarios[1]	Schirmer[2]	USWSFR[3]	MSGB4	Met rule: 2 objectives / No sSS
1	+	+	No Need	2 Objectives
2	+	-	Mandatory +	2 Objectives
			Mandatory -	No sSS
3	-	+	Mandatory +	2 Objectives
			Mandatory -	No sSS
4	-	-	No Need	No sSS

1.With assumption that the case has at least oral or ophthalmic symptom that are defined as: Oral symptoms definition – Positive response to at least one of them including: a. Having DAILY feeling of dry mouth for MORE THAN 3 MONTHS (not just feeling of dry mouth) b. FREQUENTLY drink liquids to aid in swallowing dry food (not just sometimes drink liquids to aid in swallowing dry food), c. RECURRENTLY or PERSISTENTLY swollen salivary glands in ADULT. Ophthalmic symptoms - Positive response to at least one of them including: a. Having DAILY, PERSISTENT, TROUBLESOME dry eyes for MORE THAN 3 MONTHS (not only having dry eye), b. Having a RECURRENT sensation of sand or gravel in the eyes (not only having sensation of sand or gravel in eyes), c. Using tear substitutes AT LEAST 4 TIMES A DAY (not only 3 times or less), 2. Schirmer: Schirmer test without analgesic, 3. USWSFR: Unstimulated whole salivary flow rate,4. MSGB: Minor salivary gland biopsy

[DOI: 10.47176/mjiri.40.46]

Table 12. Scenarios based decision table of ACR/EULAR criteria for primary Sjögren syndrome (pSS); combination of Schirmer test without analgesic, unstimulated whole salivary flow rate and minor salivary gland biopsy (26)

Possible scenarios	Anti Ro Ab	Schirmer[1]	USWSFR[2]	MSGB[3]		Score when has met ACR/EULAR rule / No pSS
1	+	+	+	No Need		5
2	+	+	-	No Need		4
3	+	-	+	No Need		4
4	+	-	-	Mandatory	+	6
					-	No pSS
5	-	+	+	Mandatory	+	5
					-	No pSS
6	-	+	-	Mandatory	+	4
					-	No pSS
7	-	-	+	Mandatory	+	4
					-	No pSS
8	-	-	-	No Need		No pSS

1, Schirmer: Schirmer test without analgesic, 2.USWSFR: Unstimulated whole salivary flow rate,3.MSGB:Minor salivary gland biopsy

Table 13. Scenario based decision table of ACR criteria for Sjögren syndrome (SS); combination of Anti Ro Ab, Anti la Ab, Antinuclear Ab (immunofluorescence method), Rheumatoid Factor and minor salivary gland biopsy results (27)

Possible scenarios	Serology				Score when has met ACR rule / No pSS
	Anti Ro or La Ab	FANA[1] and RF [2]	MSGB[3]		
1	+	No Need	Mandatory	+	2
				-	No SS
2	-	Mandatory +	Mandatory	+	2
				-	No SS
		-	No Need		No SS

1..FANA: Antinuclear Ab by immunofluorescence 2. RF: rheumatoid factor, 3. MSGB: Minor salivary gland biopsy

[DOI: 10.47176/mjiri.40.46]

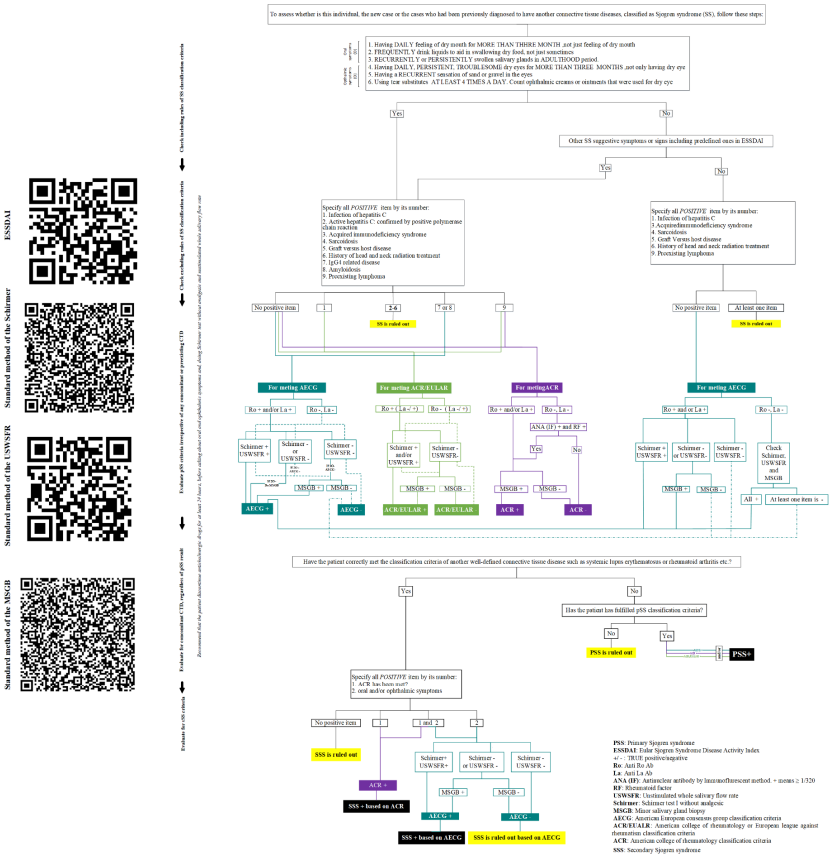


Figure 3. IRAPSS Pathway for Labeling a Case as Sjögren Syndrome, Generated Using the FEH-RAG Framework

3. Only the AECG criteria, when used to classify a case as pSS, lack a term of use and may therefore be employed for pSS screening in epidemiological studies.

4. Only the AECG and the ACR can be used for the classification of a case as sSS.

5. The ACR/EULAR is used only for the classification of cases as pSS.

6. For meeting the AECG, if either oral or ophthalmic symptoms don't exist, three of four objective tests must be positive.

7. Essentiality of using ESSDAI in clinicians' clinical routines: The ESSDAI has detailed standard definitions for SS suggestive symptoms and signs.

8. Different excluding rules of the AECG, the ACR/EULAR, and the ACR:

- Hepatitis C infection is an exclusion rule only for the AECG and the ACR not for the ACR/EULAR. In the ACR/EULAR, only if a patient has active hepatitis c, which means having a positive polymerase chain reaction (PCR), the SS will be excluded. For a case with infection of hepatitis C with negative PCR, we can only use the ACR/EULAR.

- The existence of amyloidosis or IgG4-related disease does not exclude the AECG.

- The existence of preexisting lymphoma does not exclude the rules of the ACR/EULAR and the ACR.

- If a patient is taking an anticholinergic drug, dry eye, dry mouth, Schirmer test, and unstimulated whole salivary flow rate test (USWSFR) must be checked at least one day after discontinuing it.

9. A positive HIV test does not exclude the rules of the SSCC. ONLY acquired immunodeficiency disease syndrome is.

10. Positive ANA means ANA ≥ 1/320 (immunofluorescence method). Positive ANA by the ELISA method has no value in SS classification.

11. Positive Anti-La Ab has no role in the ACR/EULAR.

12. Positivity of ANA and RF has no weight in the AECG and the ACR/EULAR.

13. In the ACR, positive ANA, if tested by the immunofluorescence method, and if it is concomitant with positive RF, is valuable for SS classification and vice versa.

14. Positive serology has no role in the classification of a case as secondary SS by the AECG.

15. In an individual who has met the AECG pSS criteria and also concurrently met another CTD criteria, it doesn't necessarily mean that this case will meet the AECG sSS criteria.

16. Irrespective of the presence of a concurrent CTD, failure to meet at least one pSS criterion does not preclude classification as sSS. According to the AECG, individuals may still be classified as having sSS if they report oral

[DOI: 10.47176/mjiri.40.46]

and/or ocular symptoms in the context of an established CTD.

17. A case without any positive serology and MSGB can be a sSS case using the AECG.

18. For classification of a case as primary or secondary SS by the ACR, positive anti-La Ab has a weight equal to positive anti-Ro Ab.

19. Seronegative patients (defined as true negative Anti Ro Ab, Anti La Ab, or negative FANA and RF) should not be forgotten as potential SS cases.

20. +/- in the IRAPSS pathway indicates TRUE positive/negative results, emphasizing that positive/negative test results should not be used for decision making. It is crucial to ensure that the results are TRUE. Each test has inherent limitations and should be performed using the standardized method. For example, a negative anti-RO antibody may be a FALSE negative due to a low titer of RO52, MSGB- may result from using an unstandardized method for performing biopsy, or may be due to reporting it inappropriately, and USWSFR – and Schirmer- may result from an unstandardized method applied to these tests.

21. To ensure standardized utilization of ESSDAI (28), the USWSFR (29), the Schirmer test I without analgesic (30), and the MSGB (31), their standard methods were provided in the IRAPSS pathway as QR codes.

This study provided a first demonstration of how the FEH-RAG framework can be practically instantiated.

Discussion

In the real world, human and financial resources are inevitably constrained, limiting our capacity to reduce adverse outcomes secondary to failed decision making (1, 32-35). Most of these errors arise from a subset of modifiable factors. Error mitigation does not require addressing all potential sources of error, but rather selectively targeting key ones (36, 37). The FEH-RAG framework focuses on the misapplication of RS as a key modifiable contributor to RS-governed decision errors. Such misapplication occurs not only due to limited expertise in using RSs, but also because of contextual pressures and cognitive load, even among users who are otherwise proficient. While the SSCC were developed for SS classification, not diagnosis, their frequent use as de facto labeling standards in practice necessitates their accurate application for mitigating the high rate of SS mislabeling all around the world (12, 13, 15).

In this POC study, the FEH-RAG provided an adapted foundation for improving the labeling of SS cases by systematically retrieving the most commonly applied SSCC. It delivered not only the SSCC decision tables and a list of their misinterpretations but also the corresponding adapted IRAPSS pathway and its decision tables for use in clinicians' daily clinical routines. These adapted outputs were designed to be readily updated once an SSCC diagnostic test becomes reliably available for routine clinical practice.

The list of the SSCC misinterpretations highlights a critical gap between the rules encoded in RSs and their real-world application. This underscores the practical value of simplified RS-based outputs, such as those provided for the SSCC. (15, 28-31).

Although the 2023 Rome international consensus on Sjögren nomenclature recommended replacing SS with Sjögren disease and sSS with Sjögren disease associated with another systemic autoimmune disease, in this POC study, we used the terms used in the SSCC (38). These terms will be updated in future versions once the corresponding changes are implemented within the SSCC.

The human derived and machine readable outputs of the FEH-RAG framework, as demonstrated for SS labeling in this POC study, enable various stakeholders, including clinicians, educators, regulatory bodies, and developers, to implement the RS logics (i.e., the SSCC logic in this study) with minimal reliance on real time consultation with overburdened experts, in both digital and non digital settings.

Artificial intelligence (AI) systems excel in automating repetitive, well-defined tasks but remain limited in high-stakes domains like medicine due to risks of opacity, hallucination, and biased recommendations (7, 19, 22, 39-45). In an era increasingly dominated by AIs (46, 47), the FEH-RAG emerges not as a competitor to these technologies but as a necessary approach to provide a foundation for mitigating their limitations through corrective and complementary mechanisms (48-50).

Data-driven AIs may require large volumes of datasets, often need to cross institutional data sharing that may be impractical, especially for rare diseases such as SS (45). These datasets may overrepresent certain populations or variables and be biased by systematic malpractices (such as those occurring in SS labeling). The biases in datasets must be approached. The adaptable FEH-RAG outputs can do this task by providing decision logs for not only auditing the quality of datasets but also ongoing oversight during a domain-specific workflow. This includes flagging deviations from RSs, enforcing guardrails, and enhancing the overall quality of reasoning (7, 22, 33, 41, 48). The datasets validated against reference standards (RSs) do more than prevent AI from amplifying human errors. They enable not only the discovery of new patterns in data but also provide the empirical basis for the timely update of the RSs. FEH-RAG outputs can be used both in data-driven AI systems and as the core decision logic in non-data-driven, rule-based AI systems.

Calibrating alarms in decision support systems (DSS), both AI based and non-AI-based, is a challenge. The FEH-RAG provides a list of real-world RS misinterpretations that developers can use as key evidence for calibrating DSS alarms.

In addition, during the 4th step ("Profiling the RS Items") of FEH-RAG, its outputs are customized to specific settings, such as the labeling of SS cases in routine clinical practice in this study. These tailored outputs can be used to build an adapted DSS (51-58).

Although the framework demonstrates the potential advantages, it has limitations, detailed below.

Occurrence of stable disagreements is theoretically improbable because FEH-RAG relies on the objective application of the RSs rather than primary evidence. However, such disagreements can potentially occur, typically reflecting differences in RS interpretation beyond textual

[DOI: 10.47176/mjiri.40.46]

ambiguities. To acknowledge this concern, methods for managing such disagreements are being refined.

In this version of FEH-RAG, only the documentation of a list of disagreements was mandatory; metadata such as the frequency, severity, temporal patterns of disagreements, and the specific disciplines responsible for each misinterpretation were not systematically documented. These characteristics will be mandatory in the next version because they empower developers, educators, regulators, and policymakers to design evidence-informed interventions.

In this POC study, systematic collection of cohort demographic and characteristics was not undertaken, as the primary objective of this FEH-RAG version was to generate outputs intended for calibration against the source RSs. Nevertheless, these data, particularly for cases in which MALs stem from 2EU misinterpretations of the SSCC (EUE) or from the source extraction errors (SEE), can provide more insight into the specific populations to whom the application of RSs is particularly vulnerable to error. Accordingly, the refined FEH-RAG framework will mandate the systematic collection and reporting of cohort demographic and characteristics specifically for MAL-associated cases attributable to EUE or SEE.

Although FEH-RAG was designed for scalability and sustainability, evaluation of these properties, in both digital and non-digital settings, is beyond the scope of this POC study and will require future research.

Conclusion

This POC study established an actionable foundation for improving SS case labeling in daily clinical practice across both digital and non-digital settings. As shown here, publishing FEH-RAG outputs while highlighting potential RS misinterpretations offers a transparent and traceable basis for augmenting decision-making in domains governed by RSs.

Acknowledgment

The authors appreciate Ahmad Reza Shain (Soft Technology Engineering) for his support.

Conflict of Interests

The authors declare that they have no competing interests.

Authors' Contributions

Supervision and critical guidance of working group activities: First author.
Conceptualization, development of the FEH-RAG framework, and POC study design: Corresponding author.
Preliminary selection of the SSCC, initial technical drafting of the SSCC elements, and generation of preliminary framework outputs: Corresponding author.
Implementation of the nine steps of the FEH-RAG framework in this POC study: All authors.
Statistical analysis: Corresponding author.
Writing – original draft: Corresponding author.
Writing – review & editing: All authors.
All authors reviewed and approved the final version of the manuscript for publication.

Ethical Considerations

This study was approved by the ethics committee of Guilan University of Medical Sciences. Its ethical code is IR.GUMS.RES.1402.430.

Funding Support

This study has no funds.

Data Availability

This data of this POC study are comprehensively included within the main text of this published article.

AI Use Statement

AI was used solely to improve grammar and readability. All AI suggestions were individually reviewed, contextually verified, and manually approved by the authors.

References

1. Newman-Toker DE, Nassery N, Schaffer AC, Yu-Moe CW, Clemens GD, Wang Z, et al. Burden of serious harms from diagnostic error in the USA. BMJ Qual Saf. 2024;33(2):109-20.
2. Winters B, Custer J, Galvagno SM Jr, Colantuoni E, Kapoor SG, Lee H, et al. Diagnostic errors in the intensive care unit: a systematic review of autopsy studies. BMJ Qual Saf. 2012;21(11):894-902.
3. Committee on Diagnostic Error in Health Care; Board on Health Care Services; Institute of Medicine; National Academies of Sciences, Engineering, and Medicine. Improving diagnosis in health care. Balogh EP, Miller BT, Ball JR, editors. Washington (DC): National Academies Press (US); 2015.
4. Panagioti M, Khan K, Keers RN, Abuzour A, Phipps D, Kontopantelis E, et al. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. BMJ. 2019;366:l4185.
5. Hanley JA, Lippman-Hand A. If nothing goes wrong, is everything all right? Interpreting zero numerators. JAMA. 1983;249(13):1743-5.
6. Dunlop M, Schwartzstein RM. Reducing diagnostic error in the intensive care unit: engaging uncertainty when teaching clinical reasoning. ATS Sch. 2020;1(4):364-71.
7. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3(11):e745-e50.
8. Vaziri S, Fakouri F, Mirzaei M, Afsharian M, Azizi M, Arab-Zozani M. Prevalence of medical errors in Iran: a systematic review and meta-analysis. BMC Health Serv Res. 2019;19(1):622.
9. World Health Organization. Global patient safety action plan 2021-2030: towards eliminating avoidable harm in health care. Geneva: World Health Organization; 2021.
10. Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. BMJ. 2000;320(7237):774-7.
11. Ahsani-Estahbanati E, Doshmangir L, Najafi B, Akbari Sari A, Sergeevich Gordeev V. Incidence rate and financial burden of medical errors and policy interventions to address them: a multi-method study protocol. Health Serv Outcomes Res Methodol. 2022;22(2):244-52.
12. Baer AN. Improving the diagnostic approach to Sjögren syndrome: a 90-year quest. Arthritis Rheumatol. 2024;76(3):348-50.
13. Beydon M, McCoy S, Nguyen Y, Sumida T, Mariette X, Seror R. Epidemiology of Sjögren syndrome. Nat Rev Rheumatol. 2024;20(3):158-69.
14. Smedby KE, Vajdic CM, Falster M, Engels EA, Martínez-Maza O, Turner J, et al. Autoimmune disorders and risk of non-Hodgkin lymphoma subtypes: a pooled analysis within the InterLymph Consortium. Blood. 2008;111(8):4029-38.
15. Fox RI, Fox CM. Sjögren syndrome: why do clinical trials fail? Rheum Dis Clin North Am. 2016;42(3):519-30.
16. Bitterman DS, Aerts H, Mak RH. Approaching autonomy in medical artificial intelligence. Lancet Digit Health. 2020;2(9):e447-e9.
17. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20(1):310.
18. Maity S, Saikia MJ. Large language models in healthcare and medical

[DOI: 10.47176/mjiri.40.46]

applications: a review. Bioengineering (Basel). 2025;12(6):631.
19. Wu Y, Lin C. Unveiling the black box: imperative for explainable AI in cardiovascular disease prevention. Lancet Reg Health West Pac. 2024;48:101145.
20. Alber DA, Yang Z, Alyakin A, Yang E, Rai S, Valliani AA, et al. Medical large language models are vulnerable to data-poisoning attacks. Nat Med. 2025;31(2):618-26.
21. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA appropriateness method user's manual. Santa Monica (CA): RAND Corporation; 2001.
22. The Lancet Digital Health. Promises and challenges of digital tools in cardiovascular care. Lancet Digit Health. 2024;6(10):e673.
23. Onakpoya IJ. Rare adverse events in clinical trials: understanding the rule of three. BMJ Evid Based Med. 2018;23(1):6.
24. Hajiabbasi A, Jamshidi A, Monfared A, Mohammadi MJ, Tangestani Nejad A, Shoaee S, et al. An early view of Iran Primary Sjogren Syndrome Registry (Guilan Province Pilot Phase), an emerging effort for a national registry establishment. Med J Islam Repub Iran. 2024;38:134.
25. Vitali C, Bombardieri S, Jonsson R, Moutsopoulos HM, Alexander EL, Carsons SE, et al. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann Rheum Dis. 2002;61(6):554-8.
26. Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: a consensus and data-driven methodology involving three international patient cohorts. Arthritis Rheumatol. 2017;69(1):35-45.
27. Shiboski SC, Shiboski CH, Criswell LA, Baer AN, Challacombe S, Lanfranchi H, et al. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. Arthritis Care Res (Hoboken). 2012;64(4):475-87.
28. Seror R, Theander E, Brun JG, Ramos-Casals M, Valim V, Dörner T, et al. Validation of EULAR primary Sjögren's syndrome disease activity (ESSDAI) and patient indexes (ESSPRI). Ann Rheum Dis. 2015;74(5):859-66.
29. Navazesh M, Kumar SKS. Measuring salivary flow: challenges and opportunities. J Am Dent Assoc. 2008;139 Suppl:35S-40S.
30. Price EJ, Tappuni AR, editors. Oxford textbook of Sjögren's syndrome. 1st ed. Oxford: Oxford University Press; 2021.
31. Fisher BA, Jonsson R, Daniels T, Bombardieri M, Brown RM, Morgan P, et al. Standardisation of labial salivary gland histopathology in clinical trials in primary Sjögren's syndrome. Ann Rheum Dis. 2017;76(7):1161-8.
32. Strain T, Brage S, Sharp SJ, Richards J, Tainio M, Ding D, et al. Use of the prevented fraction for the population to determine deaths averted by existing prevalence of physical activity: a descriptive study. Lancet Glob Health. 2020;8(7):e920-e30.
33. Edney LC, Lomas J, Karnon J, Vallejo-Torres L, Stadhouders N, Siverskog J, et al. Empirical estimates of the marginal cost of health produced by a healthcare system: methodological considerations from country-level estimates. Pharmacoeconomics. 2022;40(1):31-43.
34. Schünemann HJ, Reinap M, Piggott T, Laidmäe E, Kohler K, Põld M, et al. The ecosystem of health decision making: from fragmentation to synergy. Lancet Public Health. 2022;7(4):e378-e90.
35. Hooker EA, Mallow PJ, Oglesby MM. Characteristics and trends of emergency department visits in the United States (2010-2014). J Emerg Med. 2019;56(3):344-51.
36. Berger A, Grimm KL, Noll R, Wagner TOF. Pareto-principle in rare disease education: assessing the representation of common rare diseases in medical education and coding systems. Orphanet J Rare Dis. 2024;19(1):340.
37. Cooper L, Kang SY, Bisanzio D, Maxwell K, Rodriguez-Barraquer I, Greenhouse B, et al. Pareto rules for malaria super-spreaders and super-spreading. Nat Commun. 2019;10(1):3939.
38. Ramos-Casals M, Baer AN, Brito-Zerón MDP, Hammitt KM, Bouillot C, Retamozo S, et al. 2023 International Rome consensus for the nomenclature of Sjögren disease. Nat Rev Rheumatol. 2025;21(7):426-37.
39. For trustworthy AI, keep the human in the loop. Nat Med. 2025;31(10):3207.
40. Pantiris P, Pallis PL, Chountalas PT, Dasaklis TK. Enhancing coordination and decision making in humanitarian logistics through artificial intelligence: a grounded theory approach. Logistics. 2025;9(3):113.
41. The Lancet Digital Health. Large language models: a new chapter in digital health. Lancet Digit Health. 2024;6(1):e1.
42. Hooftman J, Dijkstra AC, Suurmeijer I, van der Bij A, Paap E, Zwaan L. Common contributing factors of diagnostic error: a retrospective analysis of 109 serious adverse event reports from Dutch hospitals. BMJ Qual Saf. 2024;33(10):642-51.
43. Pelechas E, Karagianni PG, Kaltsonoudis E. How to make AI ready for rheumatology: challenges and perspectives. Mediterr J Rheumatol. 2025;36(3):354-9.
44. Henry KE, Kornfield R, Sridharan A, Linton RC, Groh C, Wang T, et al. Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. NPJ Digit Med. 2022;5(1):97.
45. Arora A, Arora A. The promise of large language models in health care. Lancet. 2023;401(10377):641.
46. Pillutla V, Landman AB, Singh JP. Digital technology and new care pathways will redefine the cardiovascular workforce. Lancet Digit Health. 2024;6(10):e674-e6.
47. Aravazhi PS, Gunasekaran P, Benjamin NZY, Thai A, Chandrasekar KK, Kolanu ND, et al. The integration of artificial intelligence into clinical medicine: trends, challenges, and future directions. Dis Mon. 2025;71(6):101882.
48. Ntoutsi E, Fafalios P, Gadiraju U, Iosifidis V, Nejdl W, Vidal ME, et al. Bias in data-driven artificial intelligence systems: an introductory survey. Wiley Interdiscip Rev Data Min Knowl Discov. 2020;10(3):e1356.
49. Arora A, Alderman JE, Palmer J, Ganapathi S, Laws E, McCradden MD, et al. The value of standards for health datasets in artificial intelligence-based applications. Nat Med. 2023;29(11):2929-38.
50. Banerji CRS, Bhardwaj Shah A, Dabson B, Chakraborti T, Hellon V, Harbron C, et al. Clinicians must participate in the development of multimodal AI. EClinicalMedicine. 2025;84:103252.
51. Cuéllar Monreal MJ, Reig Aguado J, Font Noguera I, Poveda Andrés JL. Reduction in alert fatigue in an assisted electronic prescribing system, through the Lean Six Sigma methodology. Farm Hosp. 2017;41(1):14-30.
52. Olakotan OO, Yusof MM. Evaluating the alert appropriateness of clinical decision support systems in supporting clinical workflow. J Biomed Inform. 2020;106:103453.
53. Chien SC, Chen YL, Chien CH, Chin YP, Yoon CH, Chen CY, et al. Alerts in clinical decision support systems (CDSS): a bibliometric review and content analysis. Healthcare (Basel). 2022;10(4):601.
54. Khosa ZUN, Yusof MM. A review on CDSS alerts optimization: balancing precision and relevance. Stud Health Technol Inform. 2025;329:317-21.
55. Chen Z, Liang N, Zhang H, Li H, Yang Y, Zong X, et al. Harnessing the power of clinical decision support systems: challenges and opportunities. Open Heart. 2023;10(2):e002432.
56. Popoff B, Cabon S, Cuggia M, Bouzillé G, Clavier T. Expectations of intensive care physicians regarding an AI-based decision support system for weaning from continuous renal replacement therapy: predevelopment survey study. JMIR Med Inform. 2025;13:e63709.
57. Vallejo-Torres L, Claxton K, Edney LC, Karnon J, Lomas J, Ochalek J, et al. Challenges of calculating cost-effectiveness thresholds. Lancet Glob Health. 2023;11(10):e1508.
58. Saini V, Garcia-Armesto S, Klemperer D, Paris V, Elshaug AG, Brownlee S, et al. Drivers of poor medical care. Lancet. 2017;390(10090):178-90.

[DOI: 10.47176/mjiri.40.46]