Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Medicina
Escuela Profesional de Medicina Humana

Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025

TESIS

Para optar el Título Profesional de Médico Cirujano

AUTOR

Walter ALLCCARIMA RAMOS

ASESOR

Gadwyn SANCHEZ FELIX

Lima, Perú

2026



Reconocimiento − No Comercial − Compartir Igual − Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Allccarima W. Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025 [Tesis de pregrado]. Lima: Universidad Nacional Mayor de San Marcos, Facultad de Medicina, Escuela Profesional de Medicina Humana; 2026.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Walter Allccarima Ramos
Tipo de documento de identidad	DNI
Número de documento de identidad	76244743
URL de ORCID	https://orcid.org/0009-0004-4044-0411
Datos de asesor	
Nombres y apellidos	Gadwyn Sánchez Félix
Tipo de documento de identidad	DNI
Número de documento de identidad	08241530
URL de ORCID	https://orcid.org/0000-0003-4921-7481
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Carlos Walter Contreras Camarena
Tipo de documento	DNI
Número de documento de identidad	09042356
Miembro del jurado 1	
Nombres y apellidos	José Percy Amado Tineo
Tipo de documento	DNI
Número de documento de identidad	16782907
Miembro del jurado 2	
Nombres y apellidos	Manuel José Inostroza Fernández
Tipo de documento	DNI
Número de documento de identidad	25468233

Datos de investigación	
Línea de investigación	No aplica
Grupo de investigación	No aplica
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	Universidad Nacional Mayor de San Marcos País: Perú Departamento: Lima Distrito: La Victoria Calle: Av. Grau 755 Latitud -12.05819215 Longitud -77.0189181894387
Año o rango de años en que se realizó la investigación	Febrero 2026
URL de disciplinas OCDE	Dermatología, Enfermedades venéreas https://purl.org/pe-repo/ocde/ford#3.02.15

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, DECANA DE AMÉRICA
FACULTAD DE MEDICINA

ESCUELA PROFESIONAL DE MEDICINA HUMANA

UNMSM
Firmado digitalmente por VASQUEZ ALVA Rolando FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 06.03.2026 21:32:09 -05:00

ACTA N°0032-2026

ACTA DE SUSTENTACIÓN DE TESIS EN MODALIDAD PRESENCIAL PARA OPTAR EL TÍTULO PROFESIONAL DE MÉDICO CIRUJANO

Siendo las 9:00 horas del seis de marzo del año dos mil veintiséis, en el aula 2B del pabellón de aulas de la Facultad de Medicina, se reunió el Jurado integrado por los Doctores: Carlos Walter Contreras Camarena (Presidente), José Percy Amado Tineo (Miembro) y Manuel José Inostroza Fernández (Miembro).

Se realizó la exposición de la tesis titulada: **"Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025"**, presentado por don **WALTER ALLCCARIMA RAMOS**, para optar el Título Profesional de Médico Cirujano.

Habiendo obtenido el calificativo de:

Diecisiete (letras) ..... 17 (número)

que corresponde: Aprobado con mención honrosa

Quedando conforme con lo expuesto, se disponen a firmar la presente Acta.

Dr. José Percy Amado Tineo
Miembro

M.E. Manuel José Inostroza Fernández
Miembro

Dr. Carlos Walter Contreras Camarena
Presidente

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE MEDICINA
Escuela Profesional de Medicina Humana

Dr. HECTOR PEREYRA ZALDIVAR
Director (e)

UNIVERSIDAD NACIONAL MAYOR DE
SAN MARCOS
Universidad del Perú, Decana de América

Anexo 1

CERTIFICADO DE SIMILITUD

Yo, GADWYN SANCHEZ FELIX en mi condición de asesor acreditado con la Resolución Decanal N°000744-2026-D-FM/UNMSM de la tesis cuyo título es “Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025”, presentado por el bachiller WALTER ALLCCARIMA RAMOS para optar el Título Profesional de Médico Cirujano,

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 3% de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del título profesional.

Lima, 06 de marzo de 2026

Firma del Asesor

Nombres y Apellidos del asesor: GADWYN SANCHEZ FELIX

DNI: 08241530

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

AGRADECIMIENTOS

A la Escuela Profesional de Medicina Humana de la Universidad Nacional Mayor de San Marcos, por la formación académica recibida y por las oportunidades de aprendizaje brindadas.

A mi asesor, el Dr. Gadwyn Sánchez Félix, por su orientación, disponibilidad y rigurosidad durante el desarrollo de esta tesis, así como por sus aportes para mejorar la calidad del trabajo.

A mi familia, por su respaldo en lo académico y en lo personal, por facilitarme el tiempo y las condiciones necesarias para avanzar en esta investigación y por motivarme a culminarla.

DEDICATORIA

A mis padres, por acompañarme a lo largo de esta etapa y de toda mi formación, por su apoyo constante y por enseñarme a tomar en serio el estudio y el trabajo bien hecho.

A mis hermanos, por su paciencia y por estar presentes incluso en los momentos más demandantes de esta carrera.

A esa persona especial, quien caminó a mi lado y fue mi apoyo durante toda la travesía que significó culminar este trabajo.

ÍNDICE

RESUMEN ..... **vi**
ABSTRACT ..... **vii**
1. CAPÍTULO I: INTRODUCCIÓN **1**
1.1. Planteamiento del problema 1
1.2. Formulación del problema 4
1.3. Formulación de objetivos 5
1.4. Justificación de la investigación 5
1.5. Limitaciones del estudio 7
2. CAPÍTULO II: MARCO TEÓRICO **9**
2.1. Antecedentes de la investigación 9
2.1.1. Antecedentes internacionales 9
2.1.2. Antecedentes en Latinoamérica 11
2.1.3. Antecedentes nacionales 13
2.2. Bases teóricas 14
2.2.1. La dermatología en el contexto de la Atención Primaria 14
2.2.2. El proceso del diagnóstico visual: Del estímulo al juicio clínico 20
2.2.3. Epidemiología clínica y la medición del acierto 37
2.3. Marco conceptual 44
2.3.1. Egresado de Medicina 44
2.3.2. Precisión diagnóstica 44
2.3.3. Patrones de confusión diagnóstica 45
2.3.4. Dermatosis más frecuentes 46
2.4. Hipótesis de investigación 50
2.4.1. Hipótesis general 50
2.4.2. Hipótesis específicas 50
3. CAPÍTULO III: Metodología **51**
3.1. Tipo de investigación 51
3.2. Población 52
3.3. Muestra 52
3.4. Criterios de inclusión 53
3.5. Criterios de exclusión 53
3.6. Variables 54
3.7. Operacionalización de variables 56
3.8. Técnicas e instrumentos 58
3.8.1. Instrumento principal: Cuestionario de Precisión Diagnóstica Dermatológica mediante Imágenes Clínicas (CPDD-IC) 58
3.8.2. Validación del instrumento 60
Validez de contenido por juicio de expertos 61
Análisis de confiabilidad y propiedades psicométricas mediante estudio piloto 61
Síntesis de la validación 62
3.9. Proceso de recolección de datos 62
3.10. Análisis estadístico de los datos 63
3.11. Consideraciones éticas 64

4. CAPÍTULO IV: Resultados.....66
4.1. Características de la muestra.....66
4.2. Precisión diagnóstica.....66
4.3. Matriz de confusión multiclase.....68
4.4. Análisis comparativo de la precisión diagnóstica.....71
5. CAPÍTULO V: Discusión.....75
6. CAPÍTULO VI: Conclusiones y recomendaciones.....81
REFERENCIAS BIBLIOGRÁFICAS.....82
7. ANEXOS.....91
7.1. Anexo 1: Matriz de consistencia.....91
7.2. Anexo 2: Resolución decanal de la aprobación del proyecto de tesis.....93
7.3. Anexo 3: Consentimiento informado.....94
7.4. Anexo 4: Cuestionario de Precisión Diagnóstica Dermatológica mediante Imágenes Clínicas (CPDD-IC).....95
7.5. Anexo 5: Evaluación Likert de 5 puntos.....97
7.6. Anexo 6: Validación por juicio de expertos.....98
7.7. Anexo 7: Validación psicométrica.....110
7.8. Anexo 8: Acta de aprobación del proyecto por el Comité de Ética de Investigación.... 117

ÍNDICE DE TABLAS

Tabla 1. Características demográficas y académicas de los egresados de medicina de la UNMSM, 2025.....66
Tabla 2. Precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 según patología dermatológica específica.....67
Tabla 3. Precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 según categoría clínica de patologías dermatológicas.....67
Tabla 4. Patrones de error: principales confusiones diagnósticas de los egresados de medicina de la UNMSM del 2025.....70
Tabla 5. Patrones de error: direccionalidad de la confusión diagnóstica de los egresados de medicina de la UNMSM del 2025.....71
Tabla 6. Comparaciones post-hoc de la precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 entre categorías clínicas dermatológicas.....73
Tabla 7. Precisión diagnóstica global de los egresados de medicina de la UNMSM del 2025 según sexo.....73
Tabla 8. Precisión diagnóstica global de los egresados de medicina de la UNMSM del 2025 según sede de internado.....74

ÍNDICE DE FIGURAS

Figura 1. Precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 según patología y categoría clínica.....68
Figura 2. Matriz de confusión multiclase de los diagnósticos dermatológicos realizados por los egresados de medicina de la UNMSM del 2025.....69

RESUMEN

Objetivo: Evaluar el nivel de precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la Universidad Nacional Mayor de San Marcos (UNMSM) del 2025, describiendo su desempeño global, por patología y categorías clínicas, así como los patrones de error y diferencias según características sociodemográficas.

Metodología: Estudio observacional transversal descriptivo-comparativo. De 185 egresados, se obtuvo una muestra final de 113. Se aplicó el Cuestionario de Precisión Diagnóstica Dermatológica mediante Imágenes Clínicas (CPDD-IC) de 15 ítems. El instrumento demostró excelente validez de contenido por juicio de expertos (IVC=0,91; W de Kendall=0,64-0,69) y adecuada confiabilidad en un estudio piloto (KR-20=0,71; correlación punto-biserial promedio=0,43). Se construyó una matriz de confusión 15×15 y se aplicaron pruebas Q de Cochran, Friedman con post-hoc de Wilcoxon, y pruebas no paramétricas para comparar sexo, sede de internado y edad (α=0,05).

Resultados: La precisión diagnóstica global fue de 70,7%. Rosácea, acné vulgar y tiña corporis obtuvieron los mayores aciertos, mientras que queratosis seborreica y actínica presentaron los menores. La prueba Q de Cochran evidenció diferencias significativas entre las 15 patologías ($p<0,001$). Por categorías, la precisión fue mayor en inflamatorias e infecciosas, y menor en pigmentadas benignas y premalignas/malignas (Friedman, $p<0,001$). La matriz de confusión mostró confusiones frecuentes y asimétricas (ej. queratosis actínica por queratosis seborreica, o verruga vulgar por carcinoma basocelular). No se encontraron diferencias estadísticamente significativas según sexo, sede ni edad.

Conclusiones: Los egresados de medicina de la UNMSM del 2025 presentan un nivel de precisión diagnóstica global aceptable, pero con marcada heterogeneidad. El desempeño es adecuado en patologías inflamatorias e infecciosas, e insuficiente en lesiones pigmentadas benignas y premalignas/malignas. Estos hallazgos señalan la necesidad de reforzar el entrenamiento visual y diagnóstico diferencial del espectro pigmentado en el currículo de pregrado y evaluaciones de egreso.

Palabras clave: Dermatología; Precisión diagnóstica; Imágenes clínicas; Egresados de medicina; Educación médica; Matriz de confusión.

ABSTRACT

Objective: To evaluate the level of diagnostic accuracy in dermatology using clinical images among medical graduates of the Universidad Nacional Mayor de San Marcos (UNMSM) in 2025, describing their overall performance, by pathology and clinical categories, as well as error patterns and differences according to sociodemographic characteristics.

Methodology: A descriptive-comparative, cross-sectional observational study. From a population of 185 graduates, a final sample of 113 was obtained. The 15-item Dermatological Diagnostic Accuracy via Clinical Images Questionnaire (CPDD-IC, for its acronym in Spanish) was administered. The instrument demonstrated excellent content validity by expert judgment (CVI=0.91; Kendall's W=0.64-0.69) and adequate reliability in a pilot study (KR-20=0.71; average point-biserial correlation=0.43). A 15×15 confusion matrix was constructed, and Cochran's Q test, Friedman test with Wilcoxon post-hoc, and non-parametric tests were applied to compare sex, internship site, and age (α=0.05).

Results: The overall diagnostic accuracy was 70.7%. Rosacea, acne vulgaris, and tinea corporis obtained the highest number of correct answers, while seborrheic and actinic keratoses presented the lowest. Cochran's Q test showed significant differences among the 15 pathologies ($p<0.001$). By clinical categories, accuracy was higher for inflammatory and infectious diseases, and lower for benign pigmented and premalignant/malignant lesions (Friedman test, $p<0.001$). The confusion matrix revealed frequent and asymmetrical confusions (e.g., actinic keratosis for seborrheic keratosis, or verruca vulgaris for basal cell carcinoma). No statistically significant differences were found according to sex, internship site, or age.

Conclusions: The 2025 UNMSM medical graduates show an acceptable overall diagnostic accuracy level, but with marked heterogeneity. Performance is adequate in inflammatory and infectious pathologies, and insufficient in benign pigmented and premalignant/malignant lesions. These findings highlight the need to reinforce visual training and differential diagnosis of the pigmented spectrum in the undergraduate curriculum and graduation assessments.

Keywords: Dermatology; Diagnostic accuracy; Clinical images; Medical graduates; Medical education; Confusion matrix.

1. CAPÍTULO I: INTRODUCCIÓN

1.1. Planteamiento del problema

La importancia de la Dermatología en la Práctica Médica

La piel, como órgano más extenso del cuerpo y primera barrera frente al entorno, cumple funciones de protección, termorregulación, percepción sensorial y síntesis de compuestos esenciales, además de reflejar tanto enfermedades cutáneas primarias como manifestaciones de patologías sistémicas (1,2). Esta centralidad biológica se traduce en un elevado volumen de demanda asistencial: en múltiples sistemas de salud, las consultas por motivos dermatológicos representan entre 5% y 24% de la carga total en atención primaria, ubicando a las dermatosis entre los motivos más frecuentes de consulta médica (3,4). A nivel global, las estimaciones más recientes del GBD 2021 reportan que las enfermedades de piel generan decenas de millones de años de vida ajustados por discapacidad, consolidando a la dermatología como un problema de salud pública de alta relevancia, aunque a menudo subestimado en comparación con otras enfermedades crónicas no transmisibles (5–8).

En este contexto, la capacidad del médico general para identificar con precisión dermatosis frecuentes y distinguir lesiones benignas de entidades potencialmente malignas es un componente crítico de la calidad de atención en el primer nivel. La evidencia muestra que una evaluación dermatológica oportuna se asocia con mejores desenlaces: por ejemplo, se ha documentado que la interconsulta especializada durante la hospitalización puede acortar la estancia promedio y reducir reingresos al facilitar diagnósticos certeros de complicaciones cutáneas que de otro modo pasarían inadvertidas (9). En el caso de egresados de medicina de universidades como la UNMSM, que transitan de inmediato a escenarios de alta carga asistencial, la precisión diagnóstica en dermatología, particularmente cuando el juicio clínico se basa en la interpretación de imágenes, se vuelve un indicador clave de su preparación real para el ejercicio profesional.

El problema central: Deficiencias en competencias diagnósticas dermatológicas

A pesar de que los perfiles de egreso declaran que el médico general debe reconocer y manejar las dermatosis más prevalentes, así como identificar criterios de derivación, diversos estudios señalan brechas importantes entre estas expectativas y el desempeño efectivo (9–12). En contextos de teledermatología o de triaje visual, médicos de familia han mostrado sensibilidades inferiores al 80% para la detección de lesiones malignas, y se ha descrito que una proporción considerable de derivaciones urgentes por sospecha de cáncer cutáneo corresponde finalmente a lesiones benignas, lo que revela tanto errores de subdiagnóstico como de sobrediagnóstico con impacto directo en la eficiencia del sistema de salud (13). Si estos déficits se observan en médicos con experiencia en atención primaria, resulta especialmente pertinente cuantificar cuál es el nivel de precisión diagnóstica alcanzado por los egresados recientes cuando se les evalúa a partir de imágenes clínicas, y si dicha precisión varía según el tipo de patología o características del egresado.

La literatura en educación médica sugiere que estas deficiencias no aparecen de manera súbita en la práctica profesional, sino que se originan durante el pregrado. Evaluaciones recientes de estudiantes de medicina han documentado que el dominio de contenidos dermatológicos esenciales apenas alcanza alrededor de la mitad de los tópicos requeridos, y que las habilidades específicas de diagnóstico visual son aún más limitadas (10). Incluso entre especialistas, el diagnóstico basado en imágenes mantiene un margen de error considerable: análisis comparativos entre dermatólogos y algoritmos de inteligencia artificial han mostrado que los primeros pueden situarse alrededor de 70–75% de precisión, lo que pone de relieve la complejidad del reconocimiento de patrones y sugiere una vulnerabilidad aún mayor en médicos generales y egresados sin entrenamiento visual intensivo (14).

En el escenario sanitario peruano, estas deficiencias diagnósticas adquieren una relevancia crítica. Dada la marcada centralización de los médicos dermatólogos en hospitales de alta complejidad, el egresado de medicina, particularmente durante su Servicio Rural y Urbano Marginal de Salud (SERUMS), constituye a menudo el único filtro resolutivo en el primer nivel de atención. Si este profesional carece de

una competencia visual consolidada para identificar dermatosis frecuentes (infecciosas e inflamatorias) y discriminar lesiones de riesgo, las consecuencias directas incluyen tratamientos empíricos inadecuados, evolución desfavorable de los cuadros cutáneos y derivaciones innecesarias que saturan los escasos centros de referencia.

La Discrepancia en los Sistemas de Evaluación

Los análisis curriculares de pregrado han identificado dos factores estructurales que contribuyen a estas brechas: la escasa carga horaria dedicada a dermatología y el énfasis casi exclusivo en evaluaciones textuales, con poca o nula valoración formal de la competencia visual (10,15). En países donde se han incorporado componentes de evaluación basados en imágenes al currículo, por ejemplo, mediante estaciones OSCE o bancos de casos fotográficos estandarizados, se ha observado que la exposición sistemática a imágenes mejora la capacidad de los estudiantes para reconocer patrones clínicos, lo que respalda la necesidad de alinear enseñanza y evaluación con la naturaleza visual de la disciplina. Esta evidencia es particularmente relevante para el caso peruano, donde los egresados se enfrentan inmediatamente a escenarios clínicos en los que deben emitir juicios diagnósticos rápidos a partir de la inspección de lesiones cutáneas, a menudo sin acceso inmediato a especialistas.

La discrepancia entre la importancia clínica de la dermatología y la forma en que se evalúan las competencias en los exámenes nacionales de alto impacto constituye un elemento central del problema. En pruebas internacionales como el MIR y el USMLE, la inclusión sistemática de imágenes clínicas, frecuentemente en proporciones cercanas a una cuarta parte del total de preguntas, refleja el reconocimiento explícito de que el diagnóstico en disciplinas visuales requiere la evaluación directa de habilidades de reconocimiento de patrones (16–18). En contraste, el ENAM y el Examen Nacional de Residentado Médico en Perú mantienen un formato mayoritariamente textual, con escasa representación de casos ilustrados, lo que genera un desalineamiento entre las competencias evaluadas y las que se exigen en la práctica clínica cotidiana (19). De acuerdo con el principio de alineamiento constructivo, los estudiantes tienden a priorizar el aprendizaje de aquello que será evaluado; por ello, la ausencia de una evaluación

robusta de habilidades visuales reduce los incentivos para desarrollar sistemáticamente la competencia diagnóstica basada en imágenes (17,18).

Frente a este escenario, resulta imprescindible caracterizar de manera objetiva el nivel de precisión diagnóstica que alcanzan los egresados de medicina de la UNMSM cuando se les enfrenta a un conjunto estandarizado de imágenes clínicas de 15 patologías dermatológicas frecuentes. Esta caracterización debe abarcar: la estimación del nivel de precisión global y específico por patología; la comparación del desempeño entre categorías de enfermedades (inflamatorias, infecciosas, pigmentadas benignas y premalignas/malignas); la identificación de los pares de diagnósticos más frecuentemente confundidos mediante el análisis de la matriz de confusión; y la exploración de posibles diferencias según sexo, edad y sede de internado. Generar esta evidencia permitirá no solo dimensionar las brechas formativas que persisten al final de la carrera, sino también aportar insumos empíricos para la mejora del perfil de egreso, la revisión de los contenidos curriculares y la eventual incorporación de componentes de evaluación visual en los exámenes nacionales.

1.2. Formulación del problema

Pregunta general:

- ¿Cuál es el nivel de precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM en 2025?

Preguntas específicas:

- ¿Cuál es la precisión diagnóstica global y específica por cada una de las 15 patologías dermatológicas evaluadas mediante imágenes clínicas?
- ¿Existen diferencias en la precisión diagnóstica entre las categorías de enfermedades dermatológicas evaluadas (inflamatorias, infecciosas, pigmentadas benignas y premalignas/malignas)?
- ¿Cuáles son los pares de dermatosis más frecuentemente confundidos y qué patrones de error diagnóstico se identifican mediante el análisis de la matriz de confusión?
- ¿Existen diferencias significativas en la precisión diagnóstica según sexo, edad y sede de internado de los participantes?

1.3. Formulación de objetivos

Objetivo general:

- Determinar la precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM en 2025.

Objetivos específicos:

- Determinar la precisión diagnóstica global y por cada una de las 15 patologías dermatológicas evaluadas.
- Comparar la precisión diagnóstica entre las categorías clínicas de patologías dermatológicas evaluadas (inflamatorias, infecciosas, pigmentadas benignas y premalignas/malignas).
- Identificar los patrones de confusión diagnóstica más frecuentes y los tipos de error mediante el análisis de la matriz de confusión.
- Comparar la precisión diagnóstica según las características sociodemográficas y académicas (sexo, edad, sede de internado) de los participantes.

1.4. Justificación de la investigación

Pese a la abundante literatura internacional que muestra deficiencias relevantes en las competencias diagnósticas dermatológicas de médicos generales en formación y en ejercicio, todavía no se dispone de datos empíricos rigurosos que describan la precisión diagnóstica basada en imágenes clínicas en médicos peruanos recién egresados tras culminar su internado. Este vacío de información es especialmente preocupante si se considera que, en el contexto peruano, estos profesionales asumen rápidamente funciones asistenciales autónomas en escenarios donde la evaluación de dermatosis se realiza, en la práctica, a partir de la inspección visual sistemática de las lesiones.

El sistema de formación médica del Perú presenta, además, características estructurales que podrían modular de forma particular la consolidación de competencias dermatológicas: limitada exposición práctica a la especialidad durante

el pregrado, escasa incorporación de casos ilustrados en las evaluaciones nacionales de licenciatura y residencia (ENAM y ENARM), en contraste con exámenes internacionales donde cerca de una cuarta parte de las preguntas incluye imágenes diagnósticas (16,19), y marcada heterogeneidad en las oportunidades de entrenamiento clínico según la sede docente. En este escenario, no es posible asumir que los hallazgos de países de altos ingresos reflejan con precisión la realidad de los egresados peruanos, lo que refuerza la necesidad de generar evidencia local.

La Universidad Nacional Mayor de San Marcos, como referente nacional en la formación de médicos que se incorporan mayoritariamente al sistema público, tiene una responsabilidad académica y social directa en documentar de forma objetiva el nivel de precisión diagnóstica dermatológica de su cohorte de egreso. Esta caracterización es particularmente relevante porque un porcentaje importante de estos nuevos médicos se integrará de inmediato al Servicio Rural y Urbano Marginal de Salud (SERUMS) y a servicios de emergencia, espacios donde suelen constituir el primer, y en ocasiones único, punto de contacto para pacientes con problemas cutáneos, los cuales pueden representar hasta una cuarta parte de los motivos de consulta en el primer nivel de atención (10,11). En tales contextos, la capacidad para reconocer dermatosis prevalentes y sospechar neoplasias cutáneas se vincula de manera directa con la seguridad del paciente y con el pronóstico clínico (20).

Asimismo, el contexto latinoamericano se caracteriza por una notable variabilidad en la formación dermatológica, tanto en pregrado como en posgrado. Estudios sobre educación en dermatoscopia han evidenciado que poco más de dos tercios de los programas de residencia cuentan con un currículo formal, con grandes diferencias en horas de entrenamiento y estrategias didácticas (21). Esta heterogeneidad respalda la importancia de describir, en el caso peruano, no solo el nivel de precisión diagnóstica global de los egresados, sino también su desempeño por tipo de patología y la frecuencia de errores entre pares de diagnósticos confundidos. Además, la evidencia en educación médica ha mostrado que la evaluación orienta el aprendizaje: cuando las pruebas de alto impacto no valoran explícitamente la competencia diagnóstica visual, tanto los estudiantes como las facultades tienden a relegar el entrenamiento sistemático en esta habilidad (10,15).

En este marco, los resultados del presente estudio pueden generar aportes de alto valor en varios niveles. En primer lugar, permitirán cuantificar con precisión las brechas formativas en el perfil de egreso en dermatología, aportando insumos concretos para la revisión de contenidos, metodologías y énfasis docentes. En segundo lugar, proporcionarán una línea de base objetiva sobre el desempeño diagnóstico visual de una cohorte de egresados, frente a la cual podrán compararse futuras intervenciones curriculares o cambios en los sistemas de evaluación. En tercer lugar, ofrecerán evidencia local que respalde la necesidad de incorporar componentes visuales estandarizados en los exámenes nacionales de certificación profesional, como el ENAM, alineando mejor la evaluación con las demandas reales de la práctica clínica. En cuarto lugar, contribuirán con datos inéditos a la literatura regional sobre educación médica en dermatología, favoreciendo que las decisiones de política académica se fundamenten en indicadores propios y no solo en extrapolaciones de otros contextos. Finalmente, la metodología empleada podrá ser replicada por otras facultades del país, abriendo la puerta a estudios multicéntricos que evalúen comparativamente la calidad formativa a nivel nacional.

En última instancia, fortalecer la competencia diagnóstica dermatológica del médico egresado se espera que se traduzca en beneficios clínicos tangibles: diagnósticos más precisos, uso más racional de recursos, derivaciones oportunas al especialista y, en el caso particular de las neoplasias cutáneas, detección más temprana con impacto directo sobre la supervivencia y la calidad de vida de los pacientes (20,22).

1.5. Limitaciones del estudio

Este estudio presenta varias limitaciones que deben considerarse al interpretar los resultados.

En primer lugar, el diseño transversal solo permite valorar las competencias diagnósticas en un momento específico, sin capturar cambios en el tiempo. Aunque se planteó una evaluación censal de la cohorte 2025, la participación voluntaria introduce un posible sesgo de autoselección; por ello, las inferencias son válidas para los participantes, pero la generalización a toda la promoción debe hacerse con cautela.

En segundo lugar, la evaluación se basa en imágenes clínicas bidimensionales estáticas, que simplifican la realidad clínica, donde intervienen información táctil, distribución tridimensional de las lesiones y datos evolutivos. Además, solo se incluyeron 15 patologías, cada una representada por una única imagen, por lo que la precisión estimada no es extrapolable a dermatosis no evaluadas ni capta la variabilidad fenotípica completa.

En tercer lugar, la muestra se limita a egresados de la UNMSM, una facultad con características específicas que pueden no reflejar la realidad de otras escuelas de medicina del país; la validez externa se restringe, por tanto, a contextos formativos similares.

Finalmente, el uso de un cuestionario autoadministrado en un entorno no supervisado y la ausencia de un límite de tiempo estricto para su resolución constituyen limitaciones importantes. Estas condiciones pudieron permitir que los participantes realizaran búsquedas de información externa o consultas en internet durante la evaluación, lo que podría sobreestimar la precisión diagnóstica real. Asimismo, la carencia de una historia clínica detallada se aparta parcialmente de la práctica clínica habitual, y la variabilidad técnica entre imágenes (iluminación, contraste, saturación) podría influir adicionalmente en las respuestas.

2. CAPÍTULO II: MARCO TEÓRICO

2.1. Antecedentes de la investigación

2.1.1. Antecedentes internacionales

A nivel internacional, múltiples investigaciones han utilizado imágenes clínicas dermatológicas como herramienta para evaluar la precisión diagnóstica de distintos perfiles de profesionales de la salud, destacando que el rendimiento no solo depende del conocimiento teórico, sino también de la calidad de las imágenes, del tipo de patología evaluada y de las características del evaluador.

En teledermatología, Hassan et al. (2025) abordaron esta cuestión mediante un diseño comparativo en el que 360 pacientes consecutivos fueron fotografiados por tres grupos con diferente nivel de experiencia: usuarios sin instrucción, pacientes que recibieron una capacitación breve de cinco minutos y residentes de dermatología (23). Los autores describieron una mejora gradual de la utilidad diagnóstica conforme aumentaba la pericia técnica del fotógrafo: la concordancia diagnóstica alcanzó 79% con imágenes de calidad subóptima, 84% con calidad intermedia y 87% cuando las fotografías fueron tomadas por residentes (23). El resultado más relevante fue que la calidad diagnóstica dependía menos de la formación teórica en dermatología que del dominio técnico sobre qué elementos visuales deben ser capturados para que la imagen sea clínicamente informativa (23).

Al estratificar por tipo de enfermedad, el estudio mostró patrones de desempeño distintos según la patología. El acné vulgar mantuvo una precisión diagnóstica muy elevada (95–98%) en los tres modos de captura, lo que sugiere que sus rasgos morfológicos son lo suficientemente distintivos como para permitir un diagnóstico correcto incluso con variaciones en la calidad de la imagen (23). En contraste, las neoplasias malignas y las dermatosis papuloescamosas presentaron un rango de precisión más amplio (76–90% según el método de obtención de la fotografía), lo que indica que estas categorías requieren imágenes de mayor calidad para alcanzar un rendimiento óptimo (23). El estudio también subrayó diferencias cualitativas relacionadas con la experiencia clínica: aunque los residentes superan a los estudiantes en desempeño, tendían a ser más conservadores y menos

inclinados a formular diagnósticos definitivos que los dermatólogos asistentes, quienes mostraron mayor seguridad en sus conclusiones incluso trabajando con imágenes de calidad moderada (23). Estos hallazgos son especialmente pertinentes para el presente estudio, ya que subrayan la importancia de considerar tanto la categoría diagnóstica como el nivel de experiencia del evaluador al analizar la precisión basada en imágenes.

En relación con la equidad diagnóstica, la literatura ha puesto en evidencia que incluso los dermatólogos certificados no logran un rendimiento uniforme cuando analizan lesiones en distintos fenotipos cutáneos. Krefting et al. (2024) estudiaron a más de 100 dermatólogos alemanes en un diseño ciego en el que se presentaron 50 fotografías de diversas patologías cutáneas en dos contextos: piel clara (Fitzpatrick I–III) y piel oscura (Fitzpatrick V–VI) (24). Los resultados mostraron una marcada asimetría: en piel clara, los especialistas acertaron aproximadamente 7 de cada 10 diagnósticos, mientras que en piel oscura la tasa de aciertos se redujo a alrededor de 1 de cada 2, con una diferencia cercana a 19 puntos porcentuales ($p \leq 0.001$) (24). La autoconfianza diagnóstica siguió el mismo patrón, con mayor seguridad al evaluar piel clara, lo que sugiere la presencia de sesgos implícitos no plenamente reconocidos por los clínicos; este hallazgo es especialmente relevante para América Latina, donde es plausible que los dermatólogos reproduzcan sesgos similares cuando enfrentan fenotipos poco representados en su formación (24).

Por otra parte, el desarrollo de algoritmos de inteligencia artificial (IA) para diagnóstico dermatológico generó un notable optimismo respecto a su capacidad para igualar o incluso superar el rendimiento humano. No obstante, la revisión sistemática de Choy et al. (2023), que integró 64 estudios, mostró un escenario más matizado: muchos modelos informaron exactitudes cercanas al 90–94% para patologías comunes cuando se evaluaban en condiciones cercanas a las de entrenamiento, pero el análisis de la calidad metodológica reveló importantes limitaciones (25). Según esta revisión, el 92% de los estudios presentaba alto riesgo de sesgo (por ejemplo, en la selección de casos, el enmascaramiento y los criterios de inclusión/exclusión) y el 97% planteaba preocupaciones sustantivas sobre su aplicabilidad clínica real (25). Un problema central fue que la mayoría de algoritmos se entrenó y validó con imágenes dermatoscópicas de alta calidad obtenidas en entornos controlados, pero rara vez se probó de forma prospectiva en la atención

primaria, donde la calidad de las imágenes suele ser más variable; esta brecha entre el rendimiento en contextos idealizados y en la práctica cotidiana plantea la necesidad de interpretar con cautela los resultados de precisión y de compararlos con el desempeño de médicos generales y egresados en escenarios reales (25).

En cuanto a la formación, Ricco et al. (2024) evaluaron, mediante un diseño cuasiexperimental, si una sesión educativa de 60 minutos podría mejorar la competencia diagnóstica basada en imágenes en estudiantes de primero a cuarto año, mayoritariamente preclínicos, utilizando cinco imágenes clínicas estandarizadas de patologías frecuentes para valorar descripción morfológica, identificación diagnóstica y elección de manejo inicial (26). A pesar de una atrición considerable (sólo 37 de 65 participantes completaron el postest), en este subgrupo se observaron mejoras estadísticamente significativas en la calidad de las descripciones, en la precisión diagnóstica y en la confianza autopercibida, con incrementos marcados en las puntuaciones de descripción y en la seguridad al diagnosticar acné vulgar y otras dermatosis (26). En conjunto, estos resultados sugieren que incluso intervenciones breves pueden fortalecer de manera relevante la competencia diagnóstica visual, lo que respalda la necesidad de medir con precisión el desempeño basado en imágenes para identificar brechas específicas susceptibles de abordaje educativo (26).

2.1.2. Antecedentes en Latinoamérica

En América Latina, la producción científica orientada a medir cuantitativamente la precisión diagnóstica mediante imágenes clínicas en estudiantes, internos o egresados de medicina es escasa; la mayoría de trabajos se ha centrado en percepciones sobre la suficiencia de la formación más que en evaluaciones objetivas de competencia. Un estudio transversal realizado en Chile por Rojas et al. (2021) evaluó, mediante cuestionario autoadministrado, la percepción de 206 médicos recientemente graduados sobre la adecuación de su entrenamiento dermatológico durante el pregrado (27). Aproximadamente dos tercios (68.8%) consideraron que la formación recibida era “adecuada” para su futura práctica, aunque el estudio no incorporó pruebas estandarizadas de diagnóstico visual con imágenes clínicas (27). En el mismo trabajo se describió que las dermatosis más frecuentemente encontradas en la práctica de estos egresados

eran dermatitis atópica, micosis cutáneas e irritantes o alérgicas de contacto, lo que aporta contexto sobre el perfil de patologías que enfrentan en la atención primaria chilena (27).

Un análisis de consenso de expertos en dermatología a nivel latinoamericano publicado por Sánchez et al. (2023) que incluyó participación de dermatólogos líderes de Argentina, Brasil, Colombia, Chile, México y Uruguay realizó la identificación sistemática de barreras estructurales, clínicas y educativas que comprometían la equidad en el diagnóstico y cuidado dermatológico de poblaciones latinoamericanas. Mediante metodología de consenso modificada mediante discusión iterativa (Delphi modificado), los expertos identificaron ocho barreras principales clasificadas como: Barreras educativas: (a) conocimiento inadecuado de enfermedades dermatológicas comunes entre médicos de atención primaria, (b) déficits de capacitación en diagnóstico diferencial visual; Barreras culturales/sociales: (c) barreras lingüísticas y culturales en comunicación paciente-proveedor, (d) estigmatización social de pacientes con patología dermatológica visible; Barreras de recursos: (e) mala distribución geográfica de dermatólogos concentrados en capitales, (f) tiempo de consulta limitado en sistemas de salud sobrecargados; Barreras de políticas públicas: (g) ausencia de guías de práctica clínica locales adaptadas a epidemiología regional, (h) falta de priorización de patología dermatológica en agendas de políticas de salud pública (28).

Adicionalmente, un análisis sistemático (metaanálisis de 27 estudios) sobre confiabilidad diagnóstica en teledermatología documentó que cuando dos dermatólogos especialistas independientes evaluaban la misma fotografía clínica dermatológica (acuerdo inter-dermatólogos), concordancia diagnóstica era 71% (IC 95%: 68%-74%, evaluado mediante kappa de Cohen promedio=0.68, indicativo de acuerdo sustancial). Sin embargo, cuando se comparaban diagnósticos realizados por dermatólogos versus diagnósticos previos realizados por médicos de atención primaria no especializada sobre la misma patología (acuerdo especialista-generalista), concordancia diagnóstica era únicamente 44% (IC 95%: 38%-50%; kappa promedio=0.22, indicativo de acuerdo leve). Esta brecha importante de 27 puntos porcentuales sugiere discordancia sustancial en cómo especialistas y médicos generales reconocen e interpretan características visuales diagnósticas de patología dermatológica común (29).

Respecto a diagnósticos de bajo desempeño diagnóstico, un análisis sistemático de 64 estudios documentó que algoritmos de aprendizaje profundo mostraban exactitud diagnóstica más baja para infecciones de piel (mediana 88%, IQR 60-95; n=17), sugiriendo que esta categoría diagnóstica posee particular dificultad visual (25).

2.1.3. Antecedentes nacionales

En el contexto peruano no se han encontrado, hasta la fecha, publicaciones que evalúen de manera sistemática la precisión diagnóstica visual en dermatología mediante el uso de imágenes clínicas estandarizadas en estudiantes de medicina, egresados o médicos residentes. Este vacío se corroboró tras una búsqueda amplia en bases de datos biomédicas internacionales y regionales, así como en revistas médicas nacionales y repositorios de tesis universitarias, utilizando términos relacionados con precisión diagnóstica en dermatología y evaluación de competencias basadas en imágenes clínicas.

La ausencia de evidencia local contrasta con la producción internacional sobre el tema y representa una limitación importante para el diseño de intervenciones educativas sustentadas en datos empíricos. Sin esta información, no es posible: (a) describir con claridad las brechas formativas propias del contexto peruano, (b) comparar el desempeño de los egresados nacionales frente a referencias internacionales, (c) jerarquizar las áreas curriculares que requieren fortalecimiento prioritario y (d) definir líneas de base contra las cuales evaluar futuras reformas del currículo y de los sistemas de evaluación.

Si bien se dispone de estudios epidemiológicos sobre la frecuencia de enfermedades dermatológicas en el Perú y de análisis que evidencian retrasos en el diagnóstico de neoplasias cutáneas, estos trabajos se centran en la carga de enfermedad y en el funcionamiento del sistema de salud, pero no miden de forma directa las competencias diagnósticas del médico en formación mediante instrumentos estandarizados (30,31). En este escenario, el presente estudio se configura como el primer intento sistemático de cuantificar la precisión diagnóstica dermatológica basada en imágenes clínicas en egresados de medicina, aportando evidencia clave para orientar la mejora de la formación de pregrado en esta área.

2.2. Bases teóricas

2.2.1. La dermatología en el contexto de la Atención Primaria

Desde una perspectiva anatómica, la piel representa un órgano de proporciones variables según características antropométricas individuales. En adultos, la superficie típica oscila entre 1.5-2.0 metros cuadrados, constituyendo aproximadamente 12-15% del peso corporal total cuando se integran todas las capas. Lo significativo es que estas proporciones hacen de la piel el órgano más extenso con implicaciones funcionales críticas para homeostasis sistémica (1,2). Funciona no solamente como barrera pasiva sino como órgano activo multisistémico: protege contra agentes patógenos infecciosos (bacterias, virus, hongos) y radiación ultravioleta; regula temperatura corporal mediante mecanismos vasoactivos y sudoración; mantiene homeostasis hídrica y electrolítica; sintetiza vitamina D en respuesta a exposición solar; y media funciones sensoriales complejas (tacto, presión, dolor, temperatura) esenciales para interacción con el medio ambiente. Esta multifuncionalidad implica que patología cutánea trasciende consecuencias locales: las infecciones de piel pueden impactar en la termorregulación sistémica; la dermatitis severa compromete la barrera causando pérdida excesiva de agua y electrolitos; y síntomas dermatológicos como prurito crónico impactan profundamente calidad de vida y función sistémica (2).

El espectro de enfermedades dermatológicas es extraordinariamente amplio, con más de 2,000 entidades nosológicas distintas clasificadas en la International Classification of Diseases versión 11 (ICD-11) (5,8). Sin embargo, desde perspectiva de salud pública y carga clínica, aproximadamente 90% de consultas dermatológicas en atención primaria corresponde a menos de 30 condiciones específicas, abarcando: condiciones infecciosas frecuentes (dermatofitosis, impétigo, varicela, herpes zóster, verrugas virales); condiciones inflamatorias crónicas (dermatitis atópica, dermatitis de contacto, dermatitis seborreica, psoriasis, rosácea); neoplasias cutáneas (carcinomas basocelular y escamocelular, melanoma); condiciones pigmentarias (nevos melanocíticos, queratosis seborreicas); y trastornos funcionales (acné, urticaria, alopecia) (5,8). Esta concentración de carga en pequeño número de condiciones tiene implicaciones críticas para formación: permite que la educación dermatológica de pregrado sea

eficiente, enfocándose en diagnosticar correctamente el 90% de patología que médicos generales encontrarán, sin necesidad de profundizar en entidades rarísimas o esotéricas.

Un número relativamente pequeño de estas condiciones representa la mayor parte de la carga global de enfermedad medida en años de vida ajustados por discapacidad (DALYs). Cuando GBD 2021 analizó distribución de carga dermatológica global, cinco condiciones dominantes explicaban más del 80% de morbilidad evitable: dermatitis (combinada ~22%), acné (~17%), psoriasis (~11%), urticaria (~11%), y virales (~9%). Esta concentración de morbilidad en solo cinco condiciones tiene una implicación pedagógica crítica: los médicos generales pueden alcanzar impacto poblacional desproporcionadamente grande mejorando diagnóstico y manejo de únicamente estas cinco condiciones. (6)

El impacto de enfermedades dermatológicas trasciende significativamente la morbilidad física local, afectando múltiples dimensiones de bienestar: salud mental, relaciones sociales, desempeño académico y laboral, y autoestima. Estudios que compararon calidad de vida (medida mediante instrumentos estandarizados como SF-36 o DLQI - Dermatology Life Quality Index) documentaron que el impacto psicosocial de condiciones dermatológicas comunes, particularmente psoriasis severa, acné extenso, y dermatitis atópica moderada-severa, es comparable o superior al impacto reportado por pacientes con enfermedades sistémicas crónicas como diabetes mellitus tipo 2, hipertensión arterial controlada, y artritis reumatoide activa. (32,33).

Respecto a salud mental específicamente, un metaanálisis de 33 estudios observacionales documentó que aproximadamente 30-40% de pacientes dermatológicos que buscan atención especializada presenta comorbilidad con trastornos de salud mental diagnosticados (principalmente depresión y trastornos ansiosos), con tasas particularmente elevadas en pacientes con acné severo (hasta 50% depresión concurrente), psoriasis (40% depresión/ansiedad), vitíligo (35% depresión), y alopecia areata (25% depresión). Esta asociación bidireccional, donde enfermedad dermatológica aumenta riesgo mental y comorbilidad psiquiátrica exacerba severidad clínica dermatológica, subraya que atención dermatológica integral requiere consideración de dimensión psicosocial (34,35).

Desde la perspectiva económica, la carga de enfermedades dermatológicas es sustancial incluso en sistemas de salud de ingresos altos. En Estados Unidos, se estiman más de 85 millones de consultas ambulatorias anuales relacionadas con condiciones cutáneas, generando costo directo de atención que supera 75 mil millones de dólares USD anuales (aproximadamente 2.3% del presupuesto de salud total estadounidense), sin contabilizar costos indirectos asociados a pérdida de productividad laboral, ausentismo escolar, y discapacidad temporal o permanente que añaden estimados adicionales de 20-30 mil millones de dólares anuales. En contextos latinoamericanos con recursos más limitados, la carga económica relativa (como proporción del presupuesto de salud nacional) es típicamente mayor, aunque magnitudes absolutas sean menores, sugiriendo que en Perú el impacto económico de enfermedades dermatológicas en términos de consumo de recursos limitados es particularmente significativo (36).

Las enfermedades dermatológicas constituyen uno de los motivos de consulta más frecuentes en atención primaria a nivel mundial, aunque su prevalencia varía sustancialmente según región geográfica, características de sistema de salud, y metodología de medición empleada. La variabilidad refleja diferencias en: densidad de población con problemas cutáneos, acceso a atención dermatológica especializada (determina si casos leves son referidos o mantenidos en AP), y criterios de "consulta dermatológica" (algunas series cuentan dermatología como diagnóstico principal, otras como diagnóstico secundario).

En sistemas de salud europeos donde se ha documentado sistemáticamente frecuencia de consultas dermatológicas en medicina de familia, las proporciones varían sustancialmente entre países: desde una de cada 18 consultas en algunos centros hasta aproximadamente 1 de cada 5-6 consultas en otros, reflejando variación según densidad de población con problemas cutáneos, disponibilidad de atención dermatológica especializada, y criterios de qué se considera "consulta dermatológica" formal versus hallazgo incidental (37), con ejemplos específicos: Países Bajos 12.4% (estudio observacional de 10,000 consultas consecutivas), España 15-18% según región (38). Cuando se realiza evaluación sistemática de toda población presentada (examen físico completo de todos los pacientes independientemente del motivo de consulta), hallazgos de "69.83%-77.35% de pacientes presentan ALGUNA lesión cutánea" (37) requiere aclaración: estos

números incluyen lesiones triviales (lunares, cicatrices, lesiones solares benignas asintomáticas) que pacientes no reconocen como "problema" diagnosticable. Esto explica la brecha entre 70% "con lesión" versus 15% que "consultan por dermatología": muchas lesiones cutáneas coexisten pero no son motivo de consulta sino hallazgos incidentales. Críticamente, de las consultas dermatológicas formales que SÍ ocurren, 40-60% corresponde a diagnósticos que médicos generales podrían manejar completamente sin referencia a especialista, sugiriendo oportunidad perdida de resolver problemas en primer nivel si las competencias diagnósticas fueran adecuadas (37).

En contexto latinoamericano, la literatura documenta patrones de prevalencia similares a Europa pero con particularidades regionales y limitaciones de acceso más severas. En Chile, análisis reciente del sistema público de salud (2020-2023) reveló que consultas dermatológicas representan 2.7% del total de consultas a especialistas, aunque en atención primaria la proporción es significativamente mayor: 4.85-5.5% del total de atenciones, sugiriendo que una proporción importante de problemas dermatológicos se resuelve en primer nivel sin referencia especializada (28). En esta región, la carga de enfermedades dermatológicas es substancial: la dermatitis atópica presenta una prevalencia de hasta 25% en niños y 1-3% en adultos, representando una demanda sanitaria considerable (28). Un análisis de consenso de expertos latinoamericanos convocado en 2023 que incluyó participación de dermatólogos de Argentina, Brasil, Colombia, Chile, México y Uruguay identificó siete barreras mayores al acceso equitativo a atención dermatológica: (a) conocimiento inadecuado de patología dermatológica frecuente entre médicos de atención primaria, (b) barreras lingüísticas y culturales en comunicación paciente-proveedor, (c) estigmatización social de pacientes con dermatología visible, (d) mala distribución geográfica de especialistas concentrados en capitales, (e) ausencia de guías de práctica clínica contextualizadas a epidemiología regional, (f) tiempo de consulta limitado en sistemas de salud sobrecargados, (g) costo prohibitivo de algunos medicamentos (28).

Respecto a carga económica, estudios de 2021 documentaron disparidad sustancial en costos de tratamiento dermatológico entre países. En Perú, los pacientes enfrentaban costos anuales típicamente menores (en promedio $665 USD), mientras que México presentaba costos 6-7 veces mayores (en promedio

$4,255 USD). En Perú específicamente, esta carga representa una barrera crítica considerando que ingresos de trabajadores no calificados rondan los $400-600 USD mensuales, significando que costos dermatológicos anuales representan 1-2 meses de ingreso. (28)

Adicionalmente, un análisis sistemático reciente sobre confiabilidad diagnóstica en teledermatología encontró que el acuerdo diagnóstico general entre dermatólogos era del 71%, pero solo del 44% cuando se comparaban teledermatología con evaluaciones en atención primaria no especializada, sugiriendo una brecha sustancial en competencias diagnósticas entre especialistas y médicos generales (29). Estos hallazgos subrayan la necesidad urgente de fortalecer la formación en competencias diagnósticas dermatológicas en todos los niveles, particularmente en atención primaria donde la mayoría de pacientes latinoamericanos busca atención inicial por problemas dermatológicos.

En Perú, las enfermedades dermatológicas representan un componente importante de la carga de morbilidad poblacional, aunque estudios específicos que caractericen completamente su epidemiología en atención primaria son aún limitados. Existen condiciones asociadas a la altura como condiciones parasitarias tales como escabiosis, pediculosis y miasis, fuertemente asociadas a altitudes menores de 700 metros sobre el nivel del mar; en cambio, frente a altitudes superiores se encuentran más trastornos relacionados con radiación solar, evidenciando que la epidemiología dermatológica en Perú está estrechamente ligada a factores geográficos y ambientales específicos (30). Estos hallazgos tienen implicaciones críticas para la formación de médicos generales, requiriendo entrenamiento específico en diagnóstico y tratamiento de las dermatosis más comunes según área geográfica. Particularmente dado que, según un estudio realizado en la Amazonía peruana, aproximadamente el 94% de los pacientes que acudieron a atención primaria podrían haber recibido manejo inicial sin necesidad de derivación a especialista (30).

Complementando esta perspectiva amazónica, en áreas urbanas como Lima Metropolitana, la epidemiología de consultas dermatológicas difiere pero mantiene alta prevalencia de patología infecciosa. Durante la pandemia de COVID-19, un estudio prospectivo en dos hospitales de Lima que evaluó trabajadores de salud

encontró que el 77% desarrolló lesiones cutáneas relacionadas con uso prolongado de equipo de protección personal, siendo las lesiones más frecuentes eritema (44%), comedones (18%) y erosiones (15%), con localización predominante en puente nasal (62%) y mejillas (28%) (39). Aunque este hallazgo está contextualizado en emergencia sanitaria, subraya la relevancia de la patología dermatológica en la población peruana y la necesidad de que médicos generales reconozcan presentaciones comunes de lesiones cutáneas (39).

En contexto específico de neoplasias malignas cutáneas en Perú, la situación es particularmente preocupante y fundamenta la urgencia del presente estudio. Un análisis retrospectivo reciente que incluyó 1,136 pacientes diagnosticados con melanoma cutáneo atendidos en el Instituto Nacional de Enfermedades Neoplásicas (INEN) durante período de 10 años (2010-2019) documentó indicadores alarmantes de diagnóstico tardío: 81% de pacientes tuvo retraso diagnóstico superior a 6 meses calculado desde fecha de aparición de la lesión hasta confirmación histopatológica de melanoma; más de 60% se presentó con enfermedad en estadios avanzados (Estadio III-IV T según AJCC 8ª edición, que incluye pacientes con metástasis ganglionares o distantes) (31).

La mediana de espesor de Breslow fue 3.25 milímetros, siendo 39% de la cohorte clasificado como T4 (espesor >4 mm, profundidad máxima), comparado con proporciones de 5-15% documentadas en contextos con diagnóstico temprano establecido. Este patrón de espesor aumentado refleja que melanomas peruanos eran detectados tardíamente después de progresión vertical significativa, con implicaciones pronósticas críticas: supervivencia a 5 años para pacientes con espesor <1mm (diagnóstico temprano) es típicamente 95-98%, comparado con 30-40% para espesor >4 mm, sugiriendo que retraso diagnóstico documentado en serie peruana se traduce en pronóstico dramáticamente comprometido (31).

Las barreras atribuidas a retraso diagnóstico en análisis fueron múltiples e interconectadas: (a) bajo conocimiento público sobre signos de alarma de melanoma (únicamente 15% de población peruana en encuestas puede identificar criterios ABCDE), (b) acceso severamente limitado a atención dermatológica especializada (1.1 dermatólogos por 100,000 habitantes comparado con 4.0 recomendación OMS), (c) competencias diagnósticas subóptimas de médicos de

atención primaria que no reconocen lesiones sospechosas y por lo tanto no derivan oportunamente (31). Esta tercera barrera, competencias diagnosticadas insuficientes, constituye factor modificable mediante intervención educativa, en contraste con barreras de acceso a especialistas que son estructurales y requieren políticas de salud pública. Por lo tanto, mejorar competencias diagnósticas de médicos generales y residentes en formación tiene potencial de reducir retraso diagnóstico incluso en sistemas con limitaciones de acceso especializado.

Las disparidades en acceso a atención dermatológica en Perú son estructurales y profundas, con implicaciones críticas para diagnóstico temprano de patología maligna cutánea. La densidad nacional de dermatólogos certificados es aproximadamente 1.1 especialistas por cada 100,000 habitantes, significativamente menor que la recomendación de la Organización Mundial de la Salud de 4.0 especialistas por 100,000 (proporción nacional es 27.5% del objetivo recomendado) (40). Más crítico aún, la distribución geográfica es profundamente inequitativa: aproximadamente 72% de dermatólogos certificados ejerce en Lima Metropolitana, dejando región Andina con densidad de ~0.3 dermatólogos/100,000 y región Amazónica con densidad de ~0.05 dermatólogos/100,000 habitantes, constituyendo acceso prácticamente nulo a especialista (40). Esta desigualdad geográfica crea cascada de consecuencias: médicos generales en provincias deben manejar toda patología dermatológica sin acceso a referencia especializada, amplificando importancia crítica de competencias diagnósticas de estos médicos generales.

En síntesis, Perú enfrenta una convergencia de alta carga dermatológica, disparidades geográficas extremas en acceso a especialistas y diagnósticos tardíos de lesiones malignas con pronósticos comprometidos. Esta situación subraya la urgencia de identificar déficits específicos de formación e implementar intervenciones educativas que mejoren la preparación de futuros médicos generales para enfrentar la carga dermatológica del contexto peruano.

2.2.2. El proceso del diagnóstico visual: Del estímulo al juicio clínico

2.2.2.1. La percepción en Dermatología: De la piel a la imagen mental

La percepción visual en dermatología constituye un proceso cognitivo fundamentalmente activo y constructivo. Comienza cuando información visual de una lesión cutánea ingresa al sistema nervioso central y se integra con múltiples procesadores cerebrales paralelos, no siendo meramente "captura" pasiva de datos (41).

La investigación contemporánea en neurociencia visual y psicología cognitiva ha demostrado consistentemente que percepción visual es constructiva más que reproductiva: el cerebro genera predicciones anticipatorias sobre lo que "espera" ver basadas en: (a) conocimiento previo sobre características visuales de enfermedades, (b) experiencias clínicas acumuladas desde exposición previa a lesiones similares, y (c) contexto específico (localización anatómica de lesión, edad del paciente, síntomas reportados). Estas predicciones anticipatorias influencian cómo el observador procesa el estímulo visual presentado, fundamentando que dos médicos con diferentes bases de conocimiento pueden "ver" elementos distintos de la misma imagen, generando discordancia diagnóstica que refleja diferencias en procesos constructivos subyacentes, no diferencias en calidad de la imagen (41).

Un análisis comprehensivo publicado en 2019 que realizó revisión sistemática de literatura sobre principios cognitivos subyacentes a percepción visual en diagnóstico dermatológico identificó que precisión diagnóstica depende de cuatro pasos cognitivos interconectados: (41)

Primero, capacidad de VER CORRECTAMENTE: el clínico debe detectar características visuales presentes en la lesión (color, borde, patrón, simetría) sin ser distorsionado por expectativas previas. Segundo, capacidad de PERCIBIR ADECUADAMENTE: interpretar lo "visto" integrándose en percepción coherente que combine características múltiples (síntesis perceptiva). Tercero, capacidad de CATEGORIZAR: clasificar la percepción integrada dentro de categorías diagnósticas conocidas en base de datos mental (¿es esto un nevo? ¿un melanoma? ¿una queratosis?). Cuarto, capacidad de NOMBRAR: expresar verbalmente la categoría diagnóstica con confianza y documentar la decisión diagnóstica (41).

El análisis enfatizó que diagnóstico dermatológico se fundamenta primariamente en modalidad visual (aproximadamente 85-90% de información diagnóstica proviene de visual), con información auditiva (historia del paciente,

síntomas) y verbal (datos escritos previos) siendo importante pero secundaria (10-15% de información). Este énfasis en primacía de información visual distingue dermatología de especialidades donde la información verbal/contextual es más resaltante (ej. medicina interna donde historia es >50% de diagnóstico) (41).

El nivel fundamental de este proceso perceptivo es captación del estímulo visual, ocurriendo ya sea mediante: (a) examen clínico directo de piel con luz natural o lámparas de iluminación estándar, (b) fotografía clínica convencional de alta calidad con iluminación controlada, o (c) dermatoscopia (ampliación 10x de lesión con polarización de luz). Sin embargo, la calidad y exactitud de este estímulo visual ingresando al sistema perceptivo del observador es crítica para determinación de precisión diagnóstica subsecuente (42).

Un estudio técnico reciente de 2025 que investigó optimización de reproducción de color en imágenes dermatológicas capturadas mediante múltiples dispositivos (desde cámaras dermatoscópicas profesionales costing >$5,000 hasta teléfonos inteligentes actuales) documentó variabilidad sustancial en reproducción del color según: dispositivo utilizado, condiciones de iluminación ambiental, calibración del balance de blancos, y parámetros técnicos de apertura/velocidad de obturación. Análisis cuantitativo de coordenadas de color RGB (componentes rojo, verde, azul en escala 0-255) reveló diferencias de hasta 35-45 puntos en uno o más canales de color entre misma lesión fotografiada con diferentes dispositivos bajo condiciones similares ($p<0.001$), traduciendo a diferencias visibles a ojo humano en saturación de color, tonalidad, y percepción de multicoloración (criterio diagnóstico crítico para melanoma) (42).

Un hallazgo particularmente relevante y preocupante sobre percepción visual dermatológica es su alta susceptibilidad a sesgos cognitivos derivados de características demográficas de la población sobre la cual el clínico fue entrenado. Un estudio comparativo que evaluó a dermatólogos certificados alemanes (todos con >10 años de experiencia clínica) en diagnóstico visual de 24 fotografías de enfermedades cutáneas presentes en dos presentaciones fenotípicas—piel clara (Fitzpatrick I-III) versus piel oscura (Fitzpatrick V-VI)—encontró disparidad diagnóstica severa: dermatólogos identificaron correctamente 72.1% de diagnósticos en piel clara comparado con 52.8% en piel oscura, diferencia de 19.3

puntos porcentuales (p≤0.001; OR 2.28; IC 95%: 1.81-2.87). La confianza autopercibida en sus diagnósticos fue también significativamente menor en piel oscura (73.9% versus 61.7%, diferencia 12.2 puntos; p≤0.001). Este patrón sugiere que dermatólogos, aunque expertos, demostraron menor precisión cuando enfrentados a presentaciones fenotípicas diferentes a aquellas para las cuales fueron predominantemente entrenados. (24)

Esta observación tiene implicación directa para Perú: clínicos peruanos entrenados predominantemente con poblaciones mestizas y andinas pueden tener similar sesgo cuando evalúan lesiones en poblaciones nativas amazónicas, afro-peruanas, o de otras etnias con características fenotípicas distintas, sugiriendo que equidad diagnóstica en contexto peruano requiere exposición educativa sistemática a presentaciones fenotípicamente diversas.

La transformación de estímulos visuales bidimensionales (como fotografías) en representaciones mentales tridimensionales es otro aspecto crítico de la percepción dermatológica. Un estudio innovador de 2024 que exploró métodos para mejorar la estimación de profundidad de lesiones a partir de imágenes bidimensionales utilizó técnicas de visión computacional para generar representaciones tridimensionales holográficas de lesiones cutáneas (43). Aunque el estudio fue primariamente técnico, reveló un concepto importante: la información tridimensional crítica para diagnóstico (profundidad de invasión de melanoma, patrón de crecimiento, severidad de lesión) frecuentemente se pierde cuando se utiliza únicamente imagen bidimensional, sugiriendo que la percepción visual a partir de fotografías clínicas es intrínsecamente limitada en su capacidad de transmitir la complejidad tridimensional de lesiones cutáneas (43).

La investigación en fotografía clínica dermatológica ha documentado que aunque la fotografía clínica de alta calidad es una herramienta diagnóstica esencial, la mera presentación de fotografías está lejos de constituir por sí sola un proceso perceptivo suficiente para diagnóstico preciso. Un análisis reciente sobre fotografía clínica en dermatología enfatizó que fotografías de alta calidad que documentan forma, localización, y cambios progresivos de lesiones permiten que dermatólogos estudien lesiones en detalle (44). Sin embargo, existe variabilidad sustancial entre observadores en cómo perciben e interpretan la misma imagen fotográfica. Estudios

han demostrado que incluso dermatólogos especialistas presentan concordancia diagnóstica imperfecta cuando evalúan fotografías de la misma lesión, particularmente en categorías diagnósticas donde características visuales se solapan (por ejemplo, distinguir nevos melanocíticos benignos de melanomas tempranos, o queratosis seborreicas de carcinomas basocelulares) (24,44).

La percepción visual en dermatología también se ve profundamente influenciada por la estructura perceptiva organizada, lo que en psicología visual se denomina "Gestalt", o la capacidad del cerebro de organizar elementos visuales individuales en totalidades coherentes y significativas (41). Un clínico experimentado puede, en una mirada rápida, sintetizar múltiples características visuales (color, patrón, bordes, distribución, simetría) en una representación perceptiva integrada que dispara hipótesis diagnósticas específicas. Sin embargo, esta capacidad de síntesis visual rápida, aunque frecuentemente acertada en expertos, también puede ser fuente de errores sistemáticos en estudiantes y médicos menos experimentados que pueden formar gestalts perceptivas incompletas o incorrectas basadas en características visuales limitadas (41). Este fenómeno subraya por qué la educación en diagnóstico dermatológico basada en exposición sistemática a imágenes de alta calidad es crucial no solamente para adquirir conocimiento factual sobre enfermedades, sino para entrenar los procesos perceptivos visuales fundamentales que median entre la captación del estímulo visual y la formulación del juicio diagnóstico.

Dado que los procesos cognitivos descritos se consolidan mediante exposición repetida a ejemplos, este estudio evalúa si egresados de medicina, tras años de exposición clínica durante pregrado e internado, han consolidado patrones de diagnóstico visual correcto para 15 patologías dermatológicas comunes. La precisión documentada refleja tanto: (a) calidad de patrones diagnósticos que han internalizado (el 'prototipo' mental que reconocen), como (b) la eficacia de su educación dermatológica previa en consolidar estos patrones correctamente. Las confusiones diagnósticas identificadas revelarán qué pares de patologías comparten similitud visual lo suficiente para generar confusión genuina (similitud morfológica), versus qué confusiones reflejan déficit formativo real que requiere intervención educativa.

2.2.2.2. La psicología del reconocimiento de patrones: Del novato al experto

El reconocimiento de patrones constituye el mecanismo cognitivo fundamental que diferencia a los expertos de los novatos en diagnóstico dermatológico. Estudios que han comparado precisión diagnóstica y velocidad de respuesta según nivel de experiencia clínica consistentemente demuestran que la precisión diagnóstica visual dermatológica mejora gradualmente con experiencia acumulada. Un estudio seminal publicado en 1989 que comparó la precisión diagnóstica y tiempos de respuesta de sujetos en cinco niveles de experiencia (estudiantes de medicina preclínicos, estudiantes de medicina clínicos, residentes de medicina familiar, médicos generales, y dermatólogos especialistas) encontró hallazgos revolucionarios sobre este proceso: la precisión diagnóstica aumentó gradualmente desde 21% en estudiantes de medicina hasta 87% en dermatólogos, pero lo más significativo fue que los expertos responden significativamente más rápido que los no expertos ante lesiones típicas (45). Críticamente, cuando se presentaban diagnósticos incorrectos, los tiempos de respuesta aumentaban dramáticamente incluso en expertos, lo que contradice la noción de que errores resultan de respuestas apresuradas o negligencia, sugiriendo en cambio que errores reflejan confusión genuina entre patrones similares (45).

Este reconocimiento de patrones no es un proceso simple de "captura" visual, sino un proceso altamente sofisticado donde el clínico experimentado ha almacenado en memoria ejemplos prototípicos de enfermedades que reconoce rápidamente cuando se presentan nuevos casos (45). Un estudio reciente de 2025 que exploró cualitativamente factores que influencian el aprendizaje dermatológico en médicos en entrenamiento de medicina familiar encontró que un tema recurrente fue la mejora progresiva en confianza y competencia diagnóstica derivada de experiencia clínica acumulada (46). Los médicos en entrenamiento describen explícitamente que dermatología, como especialidad visualmente dependiente, se basa fundamentalmente en reconocimiento de patrones que una vez que han visto una presentación específica en práctica clínica (por ejemplo, "Christmas tree rash" de pityriasis rosea), saben cómo reconocerla en el futuro (46). Significativamente, estos médicos en formación también describen que antes de adquirir experiencia, no sabían "dónde buscar", y cuando comenzaban a buscar sistemáticamente,

identificaban casos que previamente habían pasado desapercibidos, sugiriendo que el aprendizaje de patrones implica no solamente reconocimiento de características visuales específicas, sino también desarrollo de estrategias de búsqueda selectiva.

El desarrollo de expertise en diagnóstico dermatológico no es simplemente una función de años de experiencia clínica. Un análisis que investigó si reconocimiento de patrones era indicador confiable de expertise diagnóstica encontró que aunque la experiencia se asociaba con expertise, esta correlación era débil, sugiriendo que años en práctica no garantizaban desempeño a nivel de experto (47). El estudio describió un concepto crítico: distinguir entre competencia (capacidad adecuada de funcionar) y expertise (desempeño a nivel superior), encontrando que únicamente a través de práctica específicamente diseñada con objetivos claros de mejora, retroalimentación informativa, y repetición deliberada de tareas representativas, se lograba avanzar de competencia a expertise (47). Un análisis comprehensivo sobre práctica deliberada en medicina que revisó 50 médicos internistas encontró que el aprendizaje en práctica médica estaba profundamente embebido en el trabajo clínico cotidiano más que siendo deliberadamente buscado (48). El estudio identificó que actividades de aprendizaje más relevantes estaban directamente relacionadas con cuidado del paciente, con consejos y retroalimentación siendo buscados cuando necesarios para proporcionar ese cuidado (48). Críticamente, los pacientes encontrados y discusiones con colegas sobre pacientes eran valorados más altamente que actividades de educación formal o actualización de literatura, sugiriendo que los patrones diagnósticos se desarrollan más efectivamente a través de aprendizaje experiencial contextualizado que a través de estudio teórico aislado.

La transición desde novato a experto en dermatología involucra cambios profundos en cómo la información visual es percibida, categorizada, y procesada. Un estudio sobre educación médica en dermatología que realizó revisión de alcance de 42 estudios de intervención educativa reconoció que mientras el razonamiento clínico dermatológico es complejo y multifacético, la mayoría de intervenciones educativas se enfocan en adquisición de conocimiento teórico más que en entrenar procesos cognitivos subyacentes de reconocimiento visual (15). Esto es particularmente problemático considerando que estudios demuestran que expertos en dermatología acceden a diagnósticos a través de reconocimiento automático

(sistema 1 de pensamiento: rápido, intuitivo, poco consciente) basado en patrones almacenados en memoria, mientras que novatos deben confiar en razonamiento analítico consciente (sistema 2: lento, deliberado, consciente) que requiere procesamiento de múltiples características (15,45). La implicación es que el entrenamiento de novatos requiere exposición sistemática y repetida a ejemplos prototípicos de enfermedades en contextos variados para que estos patrones se consoliden en memoria a largo plazo, permitiendo eventualmente acceso automático y rápido.

Un concepto emergente en educación médica es la de "práctica deliberada" aplicada específicamente al desarrollo de competencias diagnósticas visuales. Un estudio de 2021 que realizó revisión sistemática de contribuciones de práctica deliberada a formación médica basada en evidencia enfatizó que mientras el modelo original de práctica deliberada (Ericsson 2008) fue conceptualizado en contextos de adquisición de habilidades motoras en deporte y música, sus principios han sido exitosamente adaptados a educación médica para entrenar competencias diagnósticas específicas (49,50). Específicamente, cuando estos principios (identificación clara de objetivos de desempeño, retroalimentación informativa inmediata, oportunidades repetidas de práctica, y ambiente seguro para error) han sido aplicados a entrenamiento de diagnóstico dermatológico visual, estudios demuestran mejoras estadísticamente significativas en precisión diagnóstica incluso después de intervenciones educativas relativamente breves (60-90 minutos) (50).

Por ejemplo, "rapid cycle deliberate practice" combina simulación, identificación de gaps de desempeño, retroalimentación específica, y múltiples oportunidades para practicar en ambiente psicológicamente seguro donde errores son analizados y corregidos sin consecuencias punitivas (50). Cuando estas estrategias de aprendizaje han sido aplicadas a dermatología, estudios han demostrado mejoras significativas incluso después de intervenciones breves (90 minutos), con estudiantes mostrando ganancias notables en capacidad de reconocer lesiones malignas, sugiriendo que el reconocimiento de patrones dermatológicos puede ser significativamente mejorado mediante metodologías de enseñanza dirigidas específicamente a entrenar este proceso cognitivo (50).

La relevancia de estos hallazgos para egresados de medicina es crítica. Dado que los egresados se encuentran en transición entre estudiante y médico general independiente, su aprendizaje durante el internado constituye oportunidad final y crítica para consolidar patrones de diagnóstico dermatológico antes de ejercer responsabilidad clínica plena. Un estudio longitudinal que rastreó crecimiento de competencia clínica en estudiantes de medicina durante rotaciones en emergencias encontró que mejoras estadísticamente significativas en razonamiento diagnóstico ocurrían por la tercera semana de rotación, sugiriendo que existe umbral mínimo de exposición clínica necesaria para adquirir competencias diagnósticas significativas (51). Sin embargo, sin retroalimentación estructurada, exposición clínica sin más no garantiza desarrollo de patrones diagnósticos correctos—los estudiantes pueden consolidar patrones erróneos si no existe feedback correctivo explícito (48,51).

2.2.2.3. Los pilares del diagnóstico visual en Dermatología

El diagnóstico visual en dermatología se fundamenta en cuatro pilares interconectados e igualmente esenciales que permiten la transformación de datos visuales en conclusiones diagnósticas precisas. Estos pilares: la morfología, la distribución, la semiología clínica, y los criterios diagnósticos formales constituyen un framework sistemático que permite al clínico organizar la información visual percibida y compararla con el conocimiento diagnóstico previo para llegar a conclusiones válidas (52,53).

El Primer Pilar: Morfología de las Lesiones Cutáneas

La morfología se refiere a la forma, apariencia, estructura y características visuales de las lesiones cutáneas. Un análisis comprehensivo sobre diagnóstico visual en dermatología enfatiza que la descripción morfológica precisa de lesiones es el fundamento sobre el cual descansan todos los demás aspectos del diagnóstico (53). Las lesiones cutáneas se describen tradicionalmente en términos de características primarias (eritema, pápula, vesícula, pústula, nódulo, mácula, placa, etc.) que representan alteraciones estructurales básicas de la piel, y características secundarias (erosión, ulceración, costra, escamas, liquenificación, etc.) que representan evolución de lesiones primarias o daño adicional (52). Un estudio reciente sobre clasificación automática de enfermedades cutáneas mediante aprendizaje profundo que incluyó análisis de más de 100,000 imágenes

dermatológicas documentó que las características morfológicas visuales, particularmente color, textura, patrón de bordes, y homogeneidad de la lesión, son los predictores más fuertes de diagnóstico correcto, superando incluso a características clínicas contextuales como edad del paciente o localización anatómica (54). Sin embargo, el estudio también documentó que estas características morfológicas deben ser valoradas de manera integrada más que individualmente, sugiriendo que la interpretación de morfología no es simplemente un ejercicio de identificación de características aisladas sino de síntesis perceptiva de múltiples características en un patrón coherente (54).

El color de las lesiones cutáneas es particularmente crítico pero también problemático en el diagnóstico visual. Un estudio que examinó reproducibilidad interobservador de criterios clínicos ABCDE para diagnóstico de melanoma con 73 dermatólogos que evaluaron 50 lesiones pigmentadas encontró que concordancia para el criterio "C" (color/coloración) mostró valores de kappa de 0.35, siendo el más bajo entre los criterios ABCDE evaluados (55). Un análisis adicional sobre fotografía clínica en dermatología que realizó revisión sistemática de 33 estudios documentó que reproducibilidad de características visuales es altamente dependiente de estandarización de fotografía, con variaciones en iluminación, distancia focal, y calibración de color produciendo diferencias sustanciales en cómo se perciben características de color de lesiones cutáneas (56).

Este hallazgo es crítico porque características de color son diagnósticamente esenciales: por ejemplo, la presencia de multicoloración (negro, marrón, rojo, azul en una misma lesión) es criterio mayor en la regla ABCDE para sospecha de melanoma, pero la clasificación de qué constituye "multicoloración" versus variación normal de color en un nevo benigno depende fuertemente de cómo el color es percibido y categorizado por el observador (55,57). Un estudio reciente que investigó estandarización de medición de color en fotografía médica dermatológica encontró que cuando se variaba iluminación y distancia de captura fotográfica, coordenadas de color RGB (componentes de rojo, verde, azul) mostraban variaciones estadísticamente significativas ($p<0.001$ por ANOVA), demostrando que factores técnicos de fotografía pueden alterar dramáticamente la representación de color en imágenes dermatológicas (58).

Este fenómeno subraya por qué fotografía clínica de alta calidad con iluminación estandarizada y balance de color calibrado es crítica para diagnóstico preciso (56,58). Diferencias en captura fotográfica pueden alterar dramáticamente cómo características morfológicas de color son percibidas entre observadores, generando discordancia diagnóstica evitable (55,56). La estandarización de técnicas fotográficas, incluyendo posición de cámara, distancia focal consistente, iluminación controlada, y calibración de color mediante referencias de color estándar (como Color Checker cards), es fundamental para mejorar reproducibilidad de diagnósticos visuales derivados de imágenes clínicas en dermatología (56,58).

El Segundo Pilar: Distribución Topográfica y Contextual

La distribución espacial de lesiones, tanto en términos de localización anatómica como de patrón de diseminación, constituye el segundo pilar diagnóstico esencial. Ciertos diagnósticos tienen predilecciones anatómicas muy específicas que contribuyen significativamente al diagnóstico diferencial. Por ejemplo, en dermatitis atópica, los criterios diagnósticos formales de la Asociación Dermatológica Japonesa especifican que una distribución simétrica con sitios de predilección específicos (periorbital, perioral, periauricular, cuello, superficies flexoras en extremidades, tronco) con características morfológicas características (eritema, exudación, pápulas, vesículas pápulas agudas; lesiones crónicas con eritema infiltrado, liquenificación y prurito) constituye hallazgo diagnóstico crítico, con el patrón de distribución siendo tan importante como características morfológicas (52,53). Un análisis reciente sobre características clínicas y demográficas de dermatosis con patrón de distribución bimodal documentó que diferentes enfermedades tienen no solamente predilecciones anatómicas específicas sino que estas predilecciones pueden variar según edad, sexo, y otros factores, con ejemplo paradigmático siendo psoriasis: la forma infantil/adolescente tiende a tener afectación facial más prevalente, mientras que formas en adulto tienden hacia tronco y extremidades (59). La distribución también puede revelar información sobre etiopatogénesis: dermatitis de contacto ocupacional que afecta manos en distribución simulando guante sugiere exposición relacionada al trabajo, mientras que distribución similar en cara sugiere causa cosmética o medicada (52,53).

La simetría versus asimetría en la distribución de lesiones también es significativa. Un concepto fundamental es que enfermedades sistémicas e inflamatorias crónicas tienden a presentar distribución simétrica (dermatitis atópica, psoriasis, rosácea típicamente simétricas), mientras que procesos infecciosos, infestaciones, o respuestas a irritantes localizados tienden a ser asimétricas o localizadas (52,53). Sin embargo, esta generalización no es infalible: un estudio que examinó las características del rash en COVID-19 encontró que aunque lesiones tipo chilblain tenían distribución característica acral (acros, particularmente dedo de pies: 83.5% de casos), esta localización específica no era patognomónica ya que otras morfologías (livedo reticularis, acro-isquemia) tenían distribución similar pero pronóstico radicalmente diferente, indicando que distribución debe interpretarse junto con otras características para diagnóstico preciso (60).

El Tercer Pilar: Semiología Clínica y Síntomas Asociados

La semiología, el conjunto de signos y síntomas clínicos asociados a lesiones cutáneas, constituye el tercer pilar del diagnóstico dermatológico visual. Aunque el énfasis de dermatología es en características visuales, síntomas subjetivos reportados por pacientes (prurito, dolor, ardor, irritación) y signos sistémicos (fiebre, síntomas respiratorios, artralgias) proporcionan información diagnóstica crítica (52,53). Un análisis sistemático que revisó características clínicas de múltiples enfermedades cutáneas enfatizó que un abordaje diagnóstico sistemático requiere documentación de: tiempo de evolución de la erupción, síntomas asociados a lesiones, antecedentes de exposiciones relevantes, medicamentos consumidos en período relevante para etiopatogénesis, presencia de síntomas sistémicos constitucionales, e impacto en calidad de vida del paciente (53,61). Por ejemplo, presencia de prurito intenso es característica de dermatitis atópica, escabiosis, o urticaria, pero ausencia de prurito en lesión nodular pigmentada de crecimiento lento en adulto mayor sugiere nevo melanocítico benigno o queratosis seborreica más que melanoma, donde prurito o cambio de tamaño es más común (52).

La temporalidad de la erupción, cuándo comenzó, si comenzó agudamente o insidiosamente, si ha mostrado tendencia al cambio o progresión, también proporciona pistas diagnósticas. Un abordaje diagnóstico sistemático publicado recientemente que compiló recomendaciones de múltiples guías de práctica clínica

dermatológica enfatizó que timeline de aparición es crítico: erupciones urticariales que aparecen en horas y desaparecen en días sin dejar marca son características de urticaria a diferencia de las erupciones papuloescamosas que persisten semanas a meses más consistentes con dermatitis o psoriasis (52,53).

El Cuarto Pilar: Criterios Diagnósticos Formales y Integración

El cuarto pilar consiste en criterios diagnósticos formales establecidos mediante investigación sistemática, que sintetizan hallazgos morfológicos, de distribución, y semiológicos en conjuntos de criterios de mayor o menor rigor diagnóstico. Para muchas enfermedades dermatológicas comunes, criterios diagnósticos basados en evidencia han sido desarrollados y consensuados internacionalmente (52,53). Por ejemplo, los criterios de Hanifin y Rajika para dermatitis atópica requieren presencia de al menos 3 características mayores (prurito, morfología y distribución típicas, liquenificación flexural en adultos/afectación facial y extensora en niños, curso crónico o recurrente, historia personal o familiar de atopia) y al menos 3 características menores (xerosis, ictiosis, reactividad de piel a pruebas intradérmicas, elevación de IgE sérica, edad de inicio temprana, susceptibilidad a infecciones cutáneas, etc.) (52). El diagnóstico formal requiere integración de hallazgos clínicos dentro de este framework de criterios.

Para condiciones malignas, criterios visuales específicos han sido desarrollados y validados para mejorar la precisión diagnóstica. La regla ABCDE (Asimetría, Bordes irregulares, Color múltiple, Diámetro > 6mm, Evolución o cambio) fue desarrollada como herramienta mnemónica para sospecha de melanoma basada en características morfológicas observadas en malignidades, aunque con limitaciones reconocidas (52). Estudios más recientes han propuesto refinamientos: la lista de "Ugly Duckling" (lesión que se ve diferente a otras lesiones del mismo paciente), algoritmos de patrón más complejos, e incorporación de características demoscópicas (52). Sin embargo, un análisis crítico de estos criterios encontró que aunque sensibilidad para melanoma puede alcanzar 90-95% cuando aplicados por dermatólogos expertos, sensibilidad en poblaciones no expertos es significativamente menor (60-75%), sugiriendo que criterios formales requieren entrenamiento específico e interpretación experta para ser efectivos (52).

Integración de los Cuatro Pilares

El diagnóstico visual preciso en dermatología no es simplemente la evaluación de cada pilar independientemente, sino la síntesis cognitiva de todos cuatro en un patrón diagnóstico integrado. Un estudio que evaluó enfoque sistemático para diagnóstico de enfermedades cutáneas documentó que cuando clínicos aplicaban framework sistematizado que requería documentación explícita de: características morfológicas primarias y secundarias, distribución topográfica y patrón de diseminación, síntomas asociados y timeline: comparación con criterios diagnósticos formales establecidos, antes de formular diagnóstico final, la precisión diagnóstica mejoraba significativamente comparado con enfoque libre o no-estructurado (53). Este hallazgo sugiere que la enseñanza de diagnóstico dermatológico en estudiantes de medicina debe enfatizar no solamente identificación de características visuales aisladas sino entrenamiento sistemático en aplicación de framework diagnóstico que integre todos cuatro pilares de manera lógica y estructurada.

2.2.2.4. Errores diagnósticos comunes en dermatología visual

El diagnóstico visual dermatológico, a pesar de la creciente disponibilidad de herramientas tecnológicas, sigue siendo vulnerable a errores sistemáticos y cognitivos que afectan tanto a estudiantes como a clínicos experimentados. La comprensión de estos errores es crítica para contextualizar por qué las evaluaciones de competencias diagnósticas son esenciales.

Naturaleza de los Errores Diagnósticos

Los errores diagnósticos en dermatología visual se originan en múltiples niveles. Un análisis interinstitucional brasileño realizado en 2025 que evaluó concordancia diagnóstica entre médicos de atención primaria/intermedia versus dermatólogos terciarios en 113 pacientes encontró que los diagnósticos fueron modificados por dermatólogos en 53.1% de los casos derivados, con valores de kappa considerados deficientes para la mayoría de grupos diagnósticos (62). Los hallazgos más preocupantes ocurrieron en diagnósticos neoplásicos: neoplasias benignas mostraron kappa=0.0769, neoplasias malignas kappa=0.0611, y

alteraciones por exposición crónica a radiación solar kappa=0.0204, indicando concordancia prácticamente nula (62). En contraste, psoriasis fue la única condición que alcanzó concordancia sustancial (kappa=0.776), sugiriendo que ciertas patologías con características morfológicas distintivas son consistentemente reconocidas mientras que otras con presentaciones variables generan errores sistemáticos (62).

Confusiones Diagnósticas Frecuentes por Similitud Morfológica

Un hallazgo consistente en estudios es que los errores diagnósticos no son aleatorios sino que reflejan confusiones específicas entre patologías que comparten características visuales. Un análisis reciente documentó que entre las lesiones que simulan malignidad se incluyen lesiones queratósicas benignas y dermatofibromas, que pueden ser confundidas frecuentemente con lesiones malignas, requiriendo una aproximación diagnóstica refinada para su diferenciación (63). Particularmente, los nevos melanocíticos benignos pueden confundirse fácilmente con melanomas, requiriendo análisis cuidadoso de características como simetría, distribución de color, y bordes para diferenciación (63).

En relación con infecciones, un desafío particular es que tiña corporis (dermatofitosis) puede ser confundida con múltiples condiciones debido a que en piel oscura la presentación es atípica: el eritema (enrojecimiento característico) es ausente o tenue, y la demarcación de escamas periféricas es pobre, resultando en confusión con dermatitis atópica, impétigo contagiosa, dermatitis seborreica, psoriasis, lupus eritematoso sistémico, o granuloma annulare (24). Esto es particularmente significativo considerando que la tiña corporis muestra la mayor discrepancia diagnóstica entre piel clara (precisión >85%) versus piel oscura (precisión 40-60%), diferencia de 45 puntos porcentuales (24).

El Rol de Características Demográficas en Errores

Un hallazgo crítico es que la precisión diagnóstica visual varía sustancialmente según características demográficas del paciente. Un estudio que comparó precisión diagnóstica de más de 100 dermatólogos alemanes evaluando 24 fotografías de enfermedades cutáneas documentó que diagnósticos correctos fueron 72.1% en piel clara versus 52.8% en piel oscura ($p \leq 0.001$; OR 2.28) (24). El

análisis de patologías específicas reveló que diagnósticos con presentación eritematosa fueron particularmente afectados por pigmentación cutánea: tinea corporis presentó la mayor brecha (diferencia >45 puntos), seguido por psoriasis, herpes zóster, dermatitis atópica, toxicodermias (TEN), herpes simple, y acné vulgar (24). Críticamente, estas son precisamente las condiciones más frecuentes en atención primaria, sugiriendo que médicos generales entrenados predominantemente con piel clara están particularmente en riesgo de cometer errores en poblaciones con piel oscura (24).

Sesgos Cognitivos en Diagnóstico Visual

Más allá de similitud morfológica, errores diagnósticos también resultan de sesgos cognitivos inherentes al procesamiento visual. Una revisión reciente identificó múltiples fuentes de error: brecha semántica y discrepancia entre terminología dermatológica y su interpretación, frecuentemente derivada de inexperiencia o exposición clínica limitada, especialmente problemático entre clínicos no-especialistas; herramientas de evaluación de riesgo subjetivas que pueden no detectar signos tempranos de melanoma o confundir lesiones benignas como lesiones de alto riesgo; limitaciones de dispositivos diagnósticos o variaciones en calidad de dermatoscopios ya que la magnificación, e iluminación pueden impactar visualización de lesiones (63). Adicionalmente, factores externos como artefactos (suciedad, cabello), áreas de difícil visualización (cuero cabelludo, superficies mucosas), y presión aplicada durante evaluación pueden oscurecer características críticas (63).

Estudios han documentado también que incluso expertos están influenciados por sesgos externos como restricciones de tiempo, presión clínica, o contexto del paciente, factores que pueden generar tanto falsos positivos como falsos negativos (63).

Desafíos Específicos en Diagnóstico de Neoplasias Malignas

El diagnóstico de neoplasias malignas presenta desafíos particulares. Lesiones cutáneas malignas frecuentemente manifiestan síntomas patológicos mínimos en etapas tempranas, como melanoma y carcinoma basocelular que pueden permanecer asintomáticos durante períodos prolongados (64). Aunque la

regla ABCDE (Asimetría, Bordes, Color, Diámetro, Evolución) y algoritmos dermatoscópicos existen, estos requieren experiencia sustancial para aplicación correcta. Un análisis realizado en 2025 que comparó diagnósticos entre sistemas de inteligencia artificial y evaluadores expertos versus no expertos encontró que evaluadores no expertos confunden frecuentemente queratosis actínica con lesiones benignas queratósicas, carcinoma basocelular con melanoma, y melanoma con nevos benignos (63). Particularmente, el estudio documentó que mientras los sistemas de IA mostraban sensibilidad comparable a humanos, demostraban especificidad significativamente mayor, sugiriendo que humanos tendían a ser demasiado ansiosos al momento de sugerir un diagnóstico de malignidad (falsos positivos), mientras que la IA era más conservadora (63).

Variabilidad Diagnóstica por Contexto Geográfico y Entrenamiento

Factores geográficos y culturales contribuyen sustancialmente a disparidades diagnósticas. Clínicos entrenados predominantemente en ciertos contextos demográficos frecuentemente luchan con diagnóstico de lesiones en poblaciones de diferentes orígenes étnicos o tipos de piel (24). Una revisión sistemática realizada en 2025 que sintetizó aplicación de inteligencia artificial a diagnóstico dermatológico en países de bajos y medianos ingresos documentó que modelos entrenados en datasets occidentales mostraban caídas dramáticas de exactitud cuando eran testeados en poblaciones africanas: un estudio en Uganda encontró exactitud de 69.9% para pacientes caucásicos comparado con apenas 17% para pacientes de raza negra (65). Esta disparidad de >50 puntos porcentuales subraya que ausencia de diversidad en datasets de entrenamiento, tanto en educación humana como en algoritmos de AI, esto perpetúa errores sistemáticos en poblaciones subrepresentadas.

Errores por Sobre-diagnóstico versus Sub-diagnóstico

Un patrón importante documentado es que médicos de atención primaria tendían a sobrediagnosticar ciertas condiciones benignas (eczema, verrugas, infecciones) mientras simultáneamente eran subdiagnosticadas las neoplasias cutáneas (62). Este sesgo de sobrediagnóstico puede resultar en: tratamientos innecesarios, costo de salud incrementado, y ansiedad de paciente innecesaria.

Inversamente, el subdiagnóstico de neoplasias resulta en retrasos diagnósticos críticos con pronósticos potencialmente devastadores (62).

Implicaciones para Evaluación de Competencias

Estos errores documentados subrayan por qué la evaluación sistemática de precisión diagnóstica mediante imágenes clínicas en egresados de medicina es crítica. Una matriz de confusión que registra patrones específicos de error permitiría: identificar qué patologías son inherentemente más desafiantes visualmente; determinar si errores reflejan similitud morfológica inherente versus deficiencias de conocimiento específico remediables; caracterizar si sesgos demográficos existen en cohorts educativas; y orientar diseños de intervenciones educativas dirigidas a mejorar precisión diagnóstica en patologías de mayor dificultad.

Esta caracterización de errores diagnósticos comunes en dermatología visual proporciona el marco conceptual para interpretar los patrones de confusión que se identificarán en este estudio. Se anticipa que los egresados de medicina de la UNMSM, como novatos en diagnóstico dermatológico, manifestarán patrones de error consistentes con la literatura revisada:

1. Mayor dificultad en diagnósticos neoplásicos versus inflamatorios.
2. Confusiones específicas entre patologías morfológicamente similares.
3. Posible variación de precisión según fototipo cutáneo representado en imágenes.
4. Errores sistemáticos más que aleatorios, reflejando confusión genuina entre patrones similares más que respuestas al azar.

La identificación cuantitativa de estos patrones en población peruana constituirá evidencia empírica local para orientar intervenciones educativas dirigidas.

2.2.3. Epidemiología clínica y la medición del acierto

2.2.3.1. Teoría de la Prueba Diagnóstica

La evaluación de competencias diagnósticas en profesionales de salud en formación puede conceptualizarse empleando principios de epidemiología clínica, pero con otra orientación crítica: la "prueba diagnóstica" es ahora el evaluador

(egresado de medicina) cuyos diagnósticos predichos se comparan contra el diagnóstico correcto conocido, permitiendo clasificar cada respuesta como verdadera o falsa (66,67). Este cambio de perspectiva es fundamental: no se evalúa una prueba de laboratorio o imagenología, sino la capacidad diagnóstica de una persona (66,67).

Aplicados a evaluación de competencia diagnóstica en personas, los conceptos de sensibilidad y especificidad adquieren significado distinto (67,68). La sensibilidad representa la proporción de diagnósticos correctos que el evaluador logra identificar entre todos los diagnósticos correctos presentados—es decir, qué porcentaje de casos correctos efectivamente diagnosticó (67). La especificidad, por el contrario, representa la proporción de diagnósticos incorrectos que el evaluador evita seleccionar—qué porcentaje de incorrecciones logra rechazar adecuadamente (67). Un evaluador con alta sensibilidad identifica correctamente la mayoría de diagnósticos verdaderos pero podría cometer muchos falsos positivos; uno con alta especificidad evita confundirse pero podría dejar pasar casos verdaderos (67,68).

En diagnóstico dermatológico visual mediante imágenes clínicas, sensibilidad refiere a: "¿De todas las imágenes que realmente muestran patología X, cuántos porcentaje el evaluador diagnostica correctamente como X?" Especificidad refiere a: "¿De todas las imágenes que realmente NO son patología X, cuántos porcentaje el evaluador correctamente descarta como algo diferente a X?"

Para propósitos del presente estudio, la precisión (exactitud) es una métrica primaria, definida simplemente como: (# de diagnósticos correctos) / (# total de imágenes) × 100%. Sin embargo, análisis secundarios pueden explorar sensibilidad y especificidad por patología, particularmente para diagnósticos críticos como malignidades donde la sensibilidad alta es clínicamente crucial (se prefiere falso positivo = referencia innecesaria a especialista, versus falso negativo = pérdida de oportunidad diagnóstica de cáncer).

Estos conceptos, aunque importados de epidemiología clínica, no son directamente aplicables a evaluación de competencia diagnóstica porque carecen de la estructura binaria (enfermo/no enfermo) de pruebas diagnósticas clínicas tradicionales (66,68). En contexto de múltiples diagnósticos posibles (15 patologías dermatológicas), se calcula sensibilidad por categoría diagnóstica como la

proporción de casos correctamente identificados de esa patología específica entre todos los casos reales de esa patología presentados, proporcionando medida de qué tan consistentemente el evaluador reconoce esa condición en particular (68,69). La precisión global (overall accuracy, proporción de diagnósticos correctos del total de diagnósticos realizados) constituye la métrica más directa e intuitiva de competencia diagnóstica integrada del evaluador (68).

Las razones de verosimilitud también pueden adaptarse a este contexto: representan cuánto más probable es que un evaluador realice diagnóstico correcto de una categoría específica comparado con diagnóstico incorrecto (66). Estas razones permiten la ponderación probabilística del desempeño diagnóstico del evaluador en diferentes contextos clínicos (66).

La curva ROC aplicada a competencia diagnóstica de personas evalúa discriminabilidad del evaluador—su capacidad de distinguir entre categorías diagnósticas correctas e incorrectas—mediante el área bajo la curva (AUC), donde AUC cercana a 1.0 representa discriminación excelente y AUC cercana a 0.5 representa desempeño equivalente al azar (69,70). Para múltiples categorías, cada diagnóstico puede evaluarse versus todos los otros, permitiendo identificación de cuáles patologías el evaluador discrimina bien versus mal (69).

En contexto de múltiples categorías diagnósticas, precisión por categoría mide qué proporción de veces que el evaluador predijo una patología específica fue correcta; recall o sensibilidad por categoría mide qué proporción de casos reales de esa patología el evaluador identificó correctamente; y F1 score balancea ambas métricas (70,71). El F1 score es particularmente valioso para identificar patologías donde el evaluador comete errores sistemáticos o presenta desempeño inconsistente (70).

El Índice de Youden (J) (J = sensibilidad + especificidad - 1) sintetiza desempeño global del evaluador en forma de sola métrica, permitiendo comparación rápida entre evaluadores o dentro del mismo evaluador en diferentes contextos (71).

Fundamentalmente, la interpretación de todas estas métricas requiere validación externa del diagnóstico correcto de cada imagen, el gold standard, mediante consenso de expertos (66,68). La confiabilidad de toda medición de

competencia diagnóstica depende de la precisión de este gold standard; si este es cuestionable o inconsistente, las métricas resultantes perderán significado (66,68).

2.2.3.2. La matriz de confusión

La matriz de confusión es una herramienta de síntesis visual que documenta sistemáticamente cómo un evaluador clasifica múltiples instancias diagnósticas, proporcionando mucho más información que una métrica única de precisión (72,73). Estructuralmente, la matriz es una tabla bidimensional donde filas representan el diagnóstico correcto (conocido de antemano) y columnas representan el diagnóstico predicho por el evaluador; cada celda contiene la frecuencia con que esa combinación específica de diagnósticos ocurrió (73,74).

En contexto de múltiples diagnósticos posibles, la matriz de confusión para un evaluador toma la forma de una tabla n×n donde n es el número total de categorías diagnósticas distintas (en este caso, 15×15) (72,74). Los elementos en la diagonal principal de la matriz, donde diagnóstico correcto coincide exactamente con diagnóstico predicho, representan las respuestas correctas del evaluador (verdaderos positivos para cada categoría) (73). Los elementos fuera de la diagonal, donde diagnóstico correcto y predicho divergen, representan errores sistemáticos del evaluador, documentando específicamente qué confusiones diagnósticas se cometen (72–74).

La ventaja fundamental de la matriz de confusión comparada con métricas globales simples es que revela patrones específicos de error que permanecerán ocultos en una medida única de precisión (72,73). Una precisión global de 60% podría reflejar dos escenarios completamente distintos: (a) un evaluador con 60% de precisión uniforme en todas las categorías, o (b) un evaluador con precisión excelente (90%) en algunas patologías pero muy pobre (<30%) en otras (73)(2,3). Ambos escenarios producen 60% de precisión global, pero tienen implicaciones pedagógicas radicalmente diferentes para alguna intervención educativa (72,73). La matriz de confusión desentraña esta heterogeneidad, permitiendo identificación específica de fortalezas y deficiencias (73).

Para cada fila de la matriz (cada diagnóstico correcto), es posible calcular sensibilidad por categoría como la proporción de diagnósticos correctamente identificados de esa patología específica entre todos los casos reales de esa

patología presentados (73,74). Esta métrica responde: "de todas las veces que esa patología fue presentada correctamente, qué porcentaje el evaluador identificó correctamente" (73). Análogamente, para cada columna (cada diagnóstico predicho), es posible calcular precisión por categoría como la proporción de diagnósticos correctos entre todos los predichos como esa categoría, respondiendo: "de todas las veces que el evaluador predijo esa patología, qué porcentaje realmente era esa patología" (74).

La matriz de confusión permite además visualización mediante heatmaps o mapas de calor donde la intensidad del color en cada celda representa la frecuencia o proporción en esa celda (72,73). Esta visualización gráfica facilita identificación rápida de patrones: la diagonal debería mostrar colores intensos (muchos aciertos), mientras que células fuera de diagonal con colores intensos revelan confusiones frecuentes (73,74). Patrones como "filas débilmente coloreadas excepto en una columna" sugieren que esa patología raramente se diagnostica correctamente sino que se confunde consistentemente con una alternativa particular (72).

La matriz de confusión también es herramienta crítica para identificar sesgos sistemáticos en el desempeño del evaluador (73,74). Por ejemplo, si un evaluador tiene frecuentemente falsos positivos para melanoma (diagnosticando melanoma cuando en realidad es nevus), esto sugeriría sesgo hacia sobrediagnóstico de malignidad (74). En contraste, si el evaluador raramente identifica melanomas reales, esto sugeriría sesgo hacia subdiagnóstico de malignidad (74). La matriz documenta específicamente estos patrones, permitiendo caracterización detallada de no sólo "qué tan bien" el evaluador diagnóstica, sino también "cómo" diagnostica y qué tipos de errores sistemáticos comete (72,73).

Cuando múltiples evaluadores son comparados, matrices de confusión derivadas de cada uno permiten análisis comparativo directo de fortalezas y deficiencias relativas (72,73). Algunos evaluadores pueden mostrar matrices con diagonal fuerte pero elementos fuera de diagonal dispersos aleatoriamente (errores no sistemáticos), mientras otros muestran diagonal débil con elementos fuera de diagonal concentrados en patrones específicos (errores sistemáticos predecibles) (73). Esta comparación revela perfiles distintos de competencia diagnóstica que justifican intervenciones educativas diferenciadas (72).

2.2.3.3. Uso de matriz de confusión para análisis de patrones de error

El análisis detallado de la matriz de confusión constituye una estrategia analítica potente para decodificar los mecanismos subyacentes de error diagnóstico, transformando patrones de error aparentemente aleatorios en información educativa estructurada y accionable. Una matriz de confusión es una tabla de contingencia donde filas representan diagnósticos correctos (lo que realmente está en la imagen), y columnas representan diagnósticos predichos (lo que el encuestado diagnóstica). La diagonal principal contiene aciertos (donde diagnóstico predicho = diagnóstico correcto), mientras que celdas fuera de la diagonal contienen errores (confusiones diagnósticas). (72–74)

La potencia de este análisis es que permite caracterización detallada de: (72,73)

- Cuáles diagnósticos el grupo diagnostica correctamente o incorrectamente (tasas de acierto por patología)
- Cuáles pares de patologías son frecuentemente confundidas
- Si las confusiones son asimétricas (ej. confunden A como B más frecuentemente que B como A, sugiriendo dificultad direccional específica)
- Si confusiones son aleatorias (distribuidas uniformemente) vs sistemáticas (concentradas entre pares específicos, sugiriendo similitud visual genuina o déficit formativo identificable)

Por ejemplo, una matriz podría revelar que cuando imagen de verdadera psoriasis se presenta, el grupo predice correctamente psoriasis 85% de veces, pero predice incorrectamente dermatitis de contacto 10% de veces y rosácea 5% de veces. Esto revela que estos dos diagnósticos (dermatitis de contacto, rosácea) comparten suficiente similitud visual con la psoriasis para generar confusión. Análisis adicionales podrían mostrar que cuando verdadera dermatitis de contacto se presenta, el grupo confunde dermatitis de contacto como psoriasis 15% de veces (asimetría!), pero predice correctamente el 80% de veces. Esta asimetría sugiere

que grupo tiene mayor dificultad reconociendo psoriasis en presentaciones atípicas que reconociendo dermatitis de contacto típica. (72,74)

Tales patrones revelados por matriz de confusión constituyen insumo crítico para diseño de intervenciones educativas dirigidas: si grupo confunde específicamente pares A-B, entonces intervención debe enfatizar características discriminatorias entre A y B, no educación general sobre ambas patologías.

Patrones de error frecuentes en diagnóstico visual incluyen: (a) confusiones basadas en similitud morfológica, donde patologías con características visuales similares se confunden frecuentemente (ej. varios tipos de lesiones eritematosas); (b) falsos positivos de malignidad, donde las lesiones benignas se diagnostican erróneamente como malignas; (c) falsos negativos de malignidad, donde las lesiones realmente malignas se diagnostican como benignas (riesgo clínico mayor); y (d) confusiones debido a variabilidad de presentación, donde cierta patología presenta morfología atípica y no es reconocida (72–74).

La matriz de confusión también permite identificación de patologías "problemas" para el evaluador, aquellas donde sensibilidad y precisión son particularmente baja, sugiriendo brechas formativas críticas (73,74). En contraste, identifica patologías fortalezas donde el evaluador demuestra precisión excelente, informando qué competencias puede servir como punto de partida para remediación de áreas débiles (72,73).

El análisis cuantitativo de la matriz mediante métricas como F1 score por categoría (combinación de precisión y recall para cada patología) permite ranking de patologías ordenadas por dificultad relativa (74). Este ranking orienta priorización de intervenciones educativas dirigidas hacia las categorías diagnósticas con mayor déficit (72,74).

Finalmente, comparación de matrices de confusión entre cohortes distintas de evaluadores (ej. egresado de años previos versus año actual) o entre subgrupos dentro de una cohorte (ej. egresado que realizó su internado en diferentes hospitales docentes) permite identificación de brechas curriculares o variabilidad educativa que requiere atención institucional (72,73).

2.3. Marco conceptual

2.3.1. Egresado de Medicina

Un egresado de medicina es un profesional que ha completado satisfactoriamente el programa de pregrado en medicina humana, incluyendo el año de internado médico obligatorio, último período de formación pre-profesional durante el cual el estudiante realiza prácticas clínicas supervisadas por 12 meses en establecimientos de salud autorizados, completando rotaciones obligatorias por especialidades troncales (Medicina Interna, Cirugía General, Pediatría, Gineco-Obstetricia) bajo supervisión de médicos especialistas (75). Al completar el internado, egresa de la universidad y se encuentra en proceso de obtención del título profesional de Médico Cirujano (75).

Los egresados frecuentemente se incorporan al sistema de salud en contextos de atención primaria, establecimientos rurales o servicios de emergencia, donde constituyen el primer contacto médico para diversas condiciones, incluyendo problemas dermatológicos (76,77). Las condiciones dermatológicas representan una proporción sustancial de consultas en atención primaria. Estudios documentan que aproximadamente 12-24% de todas las consultas en medicina general corresponden a problemas dermatológicos (78,79). Esta elevada frecuencia subraya la importancia de que los egresados posean competencias diagnósticas dermatológicas adecuadas para reconocer patología común, identificar lesiones sospechosas que requieren derivación urgente, e iniciar manejo apropiado cuando el acceso a atención especializada no está disponible (76–78).

2.3.2. Precisión diagnóstica

Precisión diagnóstica en este contexto se refiere específicamente a la capacidad de un evaluador (egresado de medicina) de formular diagnósticos correctos cuando se presenta con información visual, en este caso, fotografías clínicas de lesiones dermatológicas. Operacionalmente, se cuantifica como el porcentaje de imágenes clínicas donde el diagnóstico formulado por el egresado coincide exactamente con el diagnóstico de referencia establecido por consenso de tres dermatólogos certificados (12,80).

En escenarios multiclase donde existe múltiples diagnósticos posibles (15 patologías dermatológicas en este estudio), la precisión diagnóstica se calcula como:

$$Precisión\ diagnóstica\ = \frac{Tasa\ de\ diagnósticos\ correcto}{Tasa\ de\ diagnósticos\ realizados}\ \times\ 100\%$$

(12,80,81).

La precisión diagnóstica difiere conceptualmente de sensibilidad o especificidad que aplican a clasificación binaria; representa en cambio una medida integrada de desempeño global diagnóstico en contexto multiclase (12,81). Una precisión diagnóstica elevada (ej. >80%) indica que el evaluador identifica correctamente la mayoría de patologías presentadas, mientras que la precisión baja (<50%) indica dificultades diagnósticas significativas (25,80).

2.3.3. Patrones de confusión diagnóstica

Patrones de confusión diagnóstica se definen operacionalmente como pares de patologías dermatológicas donde la frecuencia de confusión observada entre ellas excede la frecuencia esperada bajo diagnóstico aleatorio, documentados mediante análisis de matriz de confusión 15×15, caracterizados por tendencias predecibles en las confusiones, es decir, qué diagnósticos incorrectos predice frecuentemente cuando se presenta un diagnóstico correcto específico (12,25,81). No son errores aleatorios sino que reflejan dificultad genuina del evaluador para discriminar entre pares diagnósticos específicos que comparten similitud morfológica o visual (25,81). Un patrón de confusión se documenta mediante análisis de matriz de confusión donde se registra con exactitud cuál patología fue confundida con cuál, cuántas veces ocurrió cada confusión, y si existe asimetría (ej. frecuentemente confunde A como B, pero raramente B como A) (80,81). Estos patrones son informativos porque revelan tanto dificultades visuales inherentes a ciertos pares diagnósticos (similitud morfológica genuina) como déficits específicos de formación que pueden ser direccionados mediante intervención educativa dirigida (12,25,80).

Específicamente, un "par de confusión frecuente" se identifica como:

- Frecuencia absoluta ≥ 5 egresados cometiendo el mismo error (ej. ≥5 egresados confundieron psoriasis como dermatitis de contacto).
- Proporción ≥ 5% dentro de esa fila (ej. cuando imagen es realmente psoriasis, ≥5% de grupo predice incorrectamente dermatitis de contacto)

Para cada par de confusión identificado, se caracteriza:

- Dirección de error (¿A confundido como B más que B como A?)
- Mecanismo posible (similitud morfológica real vs. déficit formativo)
- Implicación clínica (es esta confusión clínicamente significativa? ej. confundir melanoma con nevus es crítico; confundir dos tipos de nevus es menor).

2.3.4. Dermatosis más frecuentes

Dermatosis más frecuentes en este estudio se refiere específicamente a las quince condiciones dermatológicas seleccionadas para evaluación, categoría que representa las patologías de mayor prevalencia en atención primaria, de mayor relevancia clínica y epidemiológica, y que constituyen la carga principal de enfermedad dermatológica que enfrentarán los médicos generales en el contexto peruano (12,25,81). Esta selección se fundamenta en revisión de epidemiología dermatológica internacional y latinoamericana, donde estas patologías representan 70-80% del total de consultas dermatológicas en atención primaria (12,81). La selección también equilibra representatividad de diversidad diagnóstica, incluyendo entidades inflamatorias, infecciosas, y neoplásicas, asegurando que la evaluación capture espectro amplio de competencias diagnósticas visuales requeridas en medicina general (74,80).

Para análisis y organización de la investigación, las 15 patologías dermatológicas evaluadas se categorizaron en cuatro grupos clínicos basados en fisiopatología y etiología, de acuerdo con clasificaciones estándar en dermatología reconocidas internacionalmente (7,82,83). Esta estratificación teórica no solo fundamenta la selección de patologías sino que permite análisis comparativo de

precisión diagnóstica según similitud clínica y mecanismo patogénico compartido (7,82–84).

Grupo 1 - Enfermedades Inflamatorias Crónicas (n=5 patologías):

- Dermatitis atópica (CIE-11: EA80)
- Dermatitis de contacto alérgica (CIE-11: EK00)
- Psoriasis (CIE-11: EA90)
- Rosácea (CIE-11: ED90)
- Acné vulgar (CIE-11: ED80)

Definición y justificación: Este grupo incluye trastornos cuya fisiopatología central es inflamación cutánea crónica mediada por respuesta inmune aberrante (7,82). Aunque la etiología varía (reacción atópica mediada por Th2/IL-4/IL-13 en dermatitis atópica vs. dermatitis alérgica vs. mecanismos autoinmunes Th17/IL-23 en psoriasis vs. neurovascular en rosácea vs. multifactorial folicular en acné), todos comparten el mecanismo inflamatorio como eje patogénico principal (7,82,83). La inflamación es blanco terapéutico común en estas condiciones, siendo los corticosteroides, agentes inmunomoduladores, biológicos anti-IL-17, anti-IL-23, e inhibidores JAK los tratamientos de primera línea (7,83,84). Estas patologías comparten perfiles de morbilidad similar incluyendo impacto significativo en calidad de vida del paciente y requieren manejo sistémico en formas moderadas a severas (82,83).

Grupo 2 - Enfermedades Infecciosas (n=4 patologías):

- Impétigo no bulloso (CIE-11: 1B72.1)
- Tiña corporis (CIE-11: 1F2Z)
- Molusco contagioso (CIE-11: 1E76)
- Verruga vulgar (CIE-11: 1E80)

Definición y justificación: Este grupo incluye infecciones cutáneas causadas por agentes etiológicos específicos e identificables: bacteriano (Staphylococcus aureus, Streptococcus pyogenes en impétigo), fúngico (dermatofitos como Trichophyton mentagrophytes, T. rubrum, Microsporum canis en tiña corporis; incidencia global proyectada de 1.73 mil millones de casos en 2021) y viral (Molluscipoxvirus en molusco contagioso, Papillomavirus humano en verruga

vulgar) (7,83–85). Características compartidas incluyen transmisibilidad directa (contagiosidad), presencia de agente infeccioso identificable mediante cultivo o microscopía, y respuesta terapéutica específica a antimicrobianos dirigidos (antibióticos betalactámicos, antifúngicos azólicos/equinocandinas, o terapias antivirales) (7,83,85). Estas patologías se clasifican bajo la sección de "Ciertas enfermedades infecciosas o parasitarias" en la CIE-11 (86).

Grupo 3 - Lesiones Pigmentadas Benignas (n=3 patologías):

- Nevo melanocítico común (CIE-11: 2F20.0)
- Lentigo solar CIE-11: ED61)
- Queratosis seborreica (CIE-11: 2F21)

Definición y justificación: Este grupo incluye proliferaciones benignas de melanocitos y queratinocitos epidérmicos con riesgo maligno nulo o virtualmente inexistente (7,83–85). Comparten características clínicas como lesiones pigmentadas, bien delimitadas, frecuentemente adquiridas durante vida adulta, sin potencial alguno de transformación maligna (7,83–85). Dentro de este grupo existe potencial significativo de confusión diagnóstica inter-patologías debido a similitud morfológica visual: nevus y lentigos solares pueden ser difíciles de diferenciar por pigmentación similar; queratosis seborreica presenta características demoscópicas distintivas (poros comedo-like, quistes tipo milia, patrón de giros y surcos) pero puede presentar aspecto variable confundible con lesiones pigmentadas malignas si contiene pigmentación pesada (7,84,85).

Grupo 4 - Neoplasias Malignas y Premalignas (n=3 patologías):

- Melanoma (CIE-11: 2C30)
- Carcinoma basocelular (CIE-11: 2C32)
- Queratosis actínica (CIE-11: EK90)

Definición y justificación: Este grupo incluye lesiones con potencial maligno significativo: melanoma y carcinoma basocelular constituyen malignidades cutáneas establecidas requiriendo manejo oncológico especializado; queratosis actínica representa lesión pre maligna (carcinoma in situ) con capacidad documentada de progresión a carcinoma escamocelular (tasa de progresión 5-10% de lesiones individuales durante años) (7,83–85). Aunque la histogénesis varía (melanoma surge de melanocitos vs. carcinoma basocelular y queratosis actínica surgen de

queratinocitos basales), todas requieren identificación urgente, confirmación histopatológica, y manejo oncológico especializado (83–85). La diferenciación de melanoma de lesiones benignas (especialmente nevus) representa un desafío diagnóstico crítico en educación médica, dado que aproximadamente 33% de melanomas se desarrollan de nevus preexistentes (7). Estas patologías se clasifican bajo el capítulo de Neoplasias en la CIE-11 (86).

Similitud Diagnóstica Intra-grupo e Inter-grupo:

La categorización en 5 grupos no es meramente académica sino que tiene implicaciones directas para comprensión de errores diagnósticos (63,87,88):

Confusión intra-grupo esperada: Se espera una alta confusión dentro de Grupo 1 (especialmente Dermatitis atópica ↔ Dermatitis de contacto alérgica debido a presentación eccematosa similar; menor confusión con Acné por características distintas) (63,87). Similares patrones se esperan dentro Grupo 3 (Nevus ↔ Lentigo solar ↔ Queratosis seborreica confundibles por pigmentación compartida) (87,88). Estudios sobre módulos de aprendizaje perceptual en dermatología documentan que los errores diagnósticos no se distribuyen aleatoriamente sino que concentrarse en diagnósticos de similitud morfológica, clasificándose como "errores de similitud" sistemáticos que caracterizan las dificultades intrínsecas de aprendizaje visual (87).

Confusión inter-grupo clínicamente significativa: La confusión Nevus (Grupo 3 - benigno) ↔ Melanoma (Grupo 4 - maligno) constituye confusión inter-grupo de máxima relevancia clínica, dado que diferencia entre diagnóstico benigno vs. maligno requiere intervención urgente completamente distinta (63,88,89). Aproximadamente 33% de melanomas se desarrollan a partir de nevus preexistentes, por lo que la diferenciación precisa entre lesiones pigmentadas benignas y malignas constituye imperativo diagnóstico fundamental (63).

Confusión inter-grupo clínicamente irrelevante: Confusión entre Grupo 1 (Inflamatorias) y Grupo 2 (Infecciosas) es clínicamente rara, así como confusiones entre estos grupos y Grupos 3-4, debido a diferencias clínicas sustanciales en presentación visual y características diagnósticas (63,89). El análisis comparativo de exactitud diagnóstica entre clasificadores (inteligencia artificial, dermatólogos expertos y evaluadores no expertos) evidencia que las lesiones de diferentes grupos

patogénicos presentan perfiles de error significativamente distintos, con tasas de confusión inter-grupo que permanecen virtualmente nulas (63).

Esta conceptualización permite no solo seleccionar patologías coherentes para el instrumento sino también interpretar patrones de error hallados: distinguiendo entre confusiones resultantes de similitud visual legítima (intra-grupo) versus posibles déficits formativos (inter-grupo) (63,87,88). El análisis de matriz de confusión y la clasificación de errores según tipo (similaridad, mixto, aleatorio) proporciona información relevante sobre eficacia de aprendizaje y progresión diagnóstica (87).

2.4. Hipótesis de investigación

2.4.1. Hipótesis general

El nivel de precisión diagnóstica en dermatología mediante imágenes clínicas en los egresados de medicina de la UNMSM del 2025 es de nivel intermedio o regular, reflejando las brechas en el entrenamiento diagnóstico visual reportadas en la literatura.

2.4.2. Hipótesis específicas

Hipótesis específica 1: Existen diferencias estadísticamente significativas en la precisión diagnóstica entre las distintas categorías clínicas de patologías dermatológicas evaluadas (inflamatorias, infecciosas, pigmentadas benignas y premalignas/malignas) en los egresados de medicina de la UNMSM del 2025.

Hipótesis específica 2: Existen diferencias estadísticamente significativas en la precisión diagnóstica según las características sociodemográficas y académicas (sexo, edad y sede de internado) de los egresados de medicina de la UNMSM del 2025.

3. CAPÍTULO III: Metodología

La estructura metodológica del presente estudio se presenta en la matriz de consistencia (Anexo 1), que documenta la coherencia entre pregunta de investigación, objetivos, variables, metodología y análisis.

3.1. Tipo de investigación

Clasificación General: Estudio Observacional Transversal Descriptivo-Comparativo

El estudio se clasifica como observacional, transversal y de naturaleza descriptivo-comparativa, orientado a evaluar la precisión diagnóstica en dermatología mediante imágenes clínicas estandarizadas en egresados de medicina de la UNMSM. Se considera observacional porque el investigador no aplica intervenciones educativas ni modifica la formación de los participantes, limitándose a medir el desempeño diagnóstico existente en condiciones estandarizadas. Es de corte transversal porque todos los egresados son evaluados una sola vez en un periodo acotado (febrero 2026 a marzo 2026), lo que proporciona un retrato estático de sus competencias diagnósticas en ese momento, sin permitir evaluar cambios a lo largo del tiempo.

El componente descriptivo se centra en cuantificar la precisión diagnóstica global, la precisión para cada una de las quince patologías e identificar patrones de confusión diagnóstica mediante una matriz de confusión. El componente comparativo-analítico es de carácter secundario y exploratorio, y complementa el enfoque descriptivo al evaluar si existen diferencias estadísticamente significativas en la precisión diagnóstica entre las distintas patologías y entre grupos de enfermedades dermatológicas, utilizando pruebas de hipótesis para distinguir diferencias reales de variaciones atribuibles al azar. De manera secundaria, el estudio es evaluativo, pues valora la competencia diagnóstica visual mediante un cuestionario estandarizado y objetivamente calificado, previamente validado en sus propiedades psicométricas.

3.2. Población

La población está conformada por todos los egresados de la Escuela Profesional de Medicina Humana de la Universidad Nacional Mayor de San Marcos que completaron el internado médico obligatorio y egresaron en el año 2025. Según información de la Oficina de Asuntos Académicos, el número total de egresados de medicina en el año 2025 fue de 185.

El período de evaluación se desarrolló entre febrero y marzo de 2026, cuando los egresados ya habían culminado el internado y se encontraban en proceso de trámites de egreso.

3.3. Muestra

Dado que el tamaño poblacional era reducido y accesible (N=185), se optó por un diseño censal, invitando a participar a la totalidad de los egresados elegibles. En consecuencia, no se aplicó fórmula de cálculo muestral ni muestreo probabilístico. Esta decisión se sustentó en que: (a) el estudio censal elimina el error de muestreo, (b) maximiza la capacidad analítica para evaluar subgrupos pequeños, (c) resulta administrativamente eficiente en una cohorte accesible, y (d) con tasas de respuesta esperadas de 70–80%, el tamaño de muestra efectivo es comparable al que se obtendría mediante muestreo probabilístico, pero sin introducir sesgos adicionales por selección.

Como parte de los criterios de exclusión, se descartó a los egresados que habían participado previamente en el estudio piloto del instrumento CPDD-IC, con el fin de evitar sesgo de familiaridad con las imágenes y preguntas. Esta exclusión alcanzó a 30 egresados, por lo que la población accesible efectiva fue de 155 egresados elegibles (185 totales menos 30 del piloto). De este grupo, 113 egresados completaron el cuestionario, constituyendo la muestra final analizada en el estudio.

Estrategia de maximización de tasa de participación:

Para alcanzar la tasa de respuesta objetivo, se implementaron las siguientes estrategias:

- Envío de invitación formal institucional mediante correos electrónicos oficiales de la Escuela de Medicina.

- Énfasis en la importancia académica del estudio y en su potencial contribución a la mejora curricular.
- Uso de un cuestionario digital accesible desde teléfonos inteligentes, computadoras personales o equipos de la universidad.
- Tiempo estimado de completación reducido: 20-25 minutos.
- Disponibilidad del investigador para resolución de dudas mediante contacto directo durante todo el periodo de recolección.
- Período de recolección extendido (4 semanas) para facilitar participación de egresados con diferentes horarios y compromisos.

3.4. Criterios de inclusión

Se incluyeron en el estudio aquellos participantes que cumplieron todos los siguientes criterios:

- Egresados de medicina de la UNMSM en el año 2025 que culminaron satisfactoriamente el internado médico en el período académico correspondiente.
- Haber proporcionado consentimiento informado previo a la participación.
- Contar con acceso a teléfono inteligente, computadora personal o equipo de cómputo para completar el cuestionario en formato digital.

3.5. Criterios de exclusión

Se excluyeron del estudio los egresados que cumplían cualquiera de las siguientes condiciones:

- Haber participado previamente en el desarrollo o estudio piloto del instrumento CPDD-IC, para evitar sesgo de familiaridad con el cuestionario y las imágenes.
- Auto-reporte de haber realizado una rotación electiva o pasantía formal en un servicio de Dermatología con duración superior a 4 semanas durante el internado.

- Auto-reporte de deficiencias visuales severas no corregidas (por ejemplo, daltonismo, cataratas, ceguera parcial) que pudieran impedir la adecuada percepción de características diagnósticas visuales críticas, como el color.

3.6. Variables

Variable de Estudio

- **Precisión Diagnóstica Dermatológica**

  Se define como la capacidad del egresado de medicina para identificar y nombrar correctamente una patología dermatológica al visualizar su presentación clínica en una imagen estandarizada. Esta variable se evalúa a través de tres dimensiones:

  - **Precisión Diagnóstica Global (%):** Representa el desempeño general del evaluado. Se calcula como la proporción total de diagnósticos correctos sobre el número de preguntas emitidas en el instrumento.
    - *Fórmula:* (Total de aciertos ÷ Total de ítems [15]) × 100.
    - *Rango:* 0-100%.
  - **Precisión Diagnóstica por Categoría Clínica (%):** Evalúa el desempeño según el grupo fisiopatológico de la enfermedad. Se calcula como la proporción de diagnósticos correctos dentro de cada una de las cuatro categorías evaluadas (inflamatorias, infecciosas, pigmentadas benignas, y premalignas/malignas).
    - *Fórmula:* (Aciertos en categoría k ÷ Total de ítems en categoría k) × 100.
  - **Precisión Diagnóstica por Patología Específica (%):** Identifica el nivel de dificultad de cada entidad nosológica. Se calcula como la proporción de egresados que diagnosticaron correctamente cada una de las 15 patologías individuales.
    - *Fórmula:* (Número de egresados con acierto en patología X ÷ Total de egresados evaluados) × 100.

Variable de Análisis de Error

- **Patrón de Confusión Diagnóstica**

  Se define como la tendencia sistemática de los evaluados a seleccionar un diagnóstico incorrecto específico frente a una patología presentada, revelando similitud morfológica o déficit de conocimiento. Se analiza mediante dos componentes:

  - **Matriz de Confusión Multiclase (15 x 15):** Tabla de contingencia donde las filas representan el diagnóstico correcto (Gold Standard) y las columnas el diagnóstico predicho. Permite visualizar la distribución de aciertos (diagonal principal) y errores (fuera de la diagonal).
  - **Pares de Confusión Frecuentes:** Identificación de los pares específicos "Diagnóstico Real vs. Diagnóstico Errado" que presentan una frecuencia de ocurrencia superior al azar, permitiendo distinguir entre errores aleatorios y confusiones sistemáticas por similitud visual.

Variables Sociodemográficas y Académicas (De Caracterización)

- **Edad:** Tiempo transcurrido en años cumplidos desde el nacimiento hasta el momento de la evaluación.
  - *Tipo:* Variable cuantitativa de razón (continua).
- **Sexo Biológico:** Características biológicas que diferencian a los participantes.
  - *Tipo:* Variable cualitativa nominal dicotómica (Categorías: Masculino, Femenino).
- **Sede de Internado:** Institución prestadora de servicios de salud donde el egresado realizó sus prácticas pre-profesionales durante el último año académico.
  - *Tipo:* Variable cualitativa nominal politómica (Categorías: EsSalud, MINSA, Fuerzas Armadas/Policiales, Clínicas Privadas).

3.7. Operacionalización de variables

Variable principal

Variable	Definición Conceptual	Definición Operacional	Dimensiones	Indicadores	Escala de Medición	Instrumento
Precisión Diagnóstica Dermatológica	Capacidad del evaluado para identificar y nombrar correctamente una patología dermatológica al visualizar su presentación clínica en una imagen estandarizada.	Se medirá mediante el puntaje obtenido en el instrumento CPDD-IC, calculando la proporción de respuestas correctas sobre el total de ítems evaluados, tanto a nivel global como por categoría clínica y por patología específica.	Global	- Porcentaje de aciertos totales (Aciertos/15 * 100).	Razón (0-100%)	CPDD-IC
			Por Categoría Clínica	- % Acierto en Inflamatorias (5 ítems) - % Acierto en Infecciosas (4 ítems) - % Acierto en Pigmentadas Benignas (3 ítems) - % Acierto en Malignas/Premalignas (3 ítems)	Razón (0-100%)	CPDD-IC
			Por Patología Específica	- Tasa de acierto para cada una de las 15 patologías individuales (Frecuencia de acierto por ítem).	Razón (0-100%)	CPDD-IC

Variable de análisis de error

Variable	Definición Conceptual	Definición Operacional	Dimensiones	Indicadores	Escala de Medición	Instrumento
Patrón de Confusión Diagnóstica	Tendencia sistemática a seleccionar un diagnóstico incorrecto específico frente a una patología presentada,	Se obtendrá mediante la construcción de una matriz de confusión de 15×15	Error Diagnóstico	- Pares de confusión más frecuentes (patología real y diagnóstico seleccionado). - Frecuencia absoluta del error por par de confusión.	Nominal	CPDD-IC

	revelando similitud morfológica percibida o déficit de conocimiento.	(Diagnóstico real vs. diagnóstico predicho), identificando las celdas fuera de la diagonal principal con mayor frecuencia y analizando su tipo y direccionalidad.		- Proporción del error respecto al total de respuestas de la patología real. - Tipo de error (intra-categoría clínica vs. inter-categoría clínica).		

Variables sociodemográficas y académicas

Variable	Definición Conceptual	Definición Operacional	Indicadores	Escala de Medición	Instrumento
Edad	Tiempo transcurrido desde el nacimiento hasta la fecha de la evaluación.	Autorreporte del participante indicando años cumplidos.	- Años cumplidos (Número entero).	Razón	CPDD-IC
Sexo Biológico	Características biológicas y fisiológicas que definen a hombres y mujeres.	Selección de la categoría correspondiente en la ficha de recolección.	1. Masculino 2. Femenino	Nominal (Dicotómica)	CPDD-IC
Sede de Internado	Institución prestadora de servicios de salud donde el egresado realizó sus prácticas pre-profesionales durante el último año.	Selección de la categoría institucional correspondiente.	1. MINSA 2. EsSalud 3. Fuerzas Armadas / Policiales 4. Clínicas Privadas	Nominal	CPDD-IC

3.8. Técnicas e instrumentos

3.8.1. **Instrumento principal:** Cuestionario de Precisión Diagnóstica Dermatológica mediante Imágenes Clínicas (CPDD-IC).

El instrumento consiste en un cuestionario digital autoadministrado, desarrollado específicamente para esta investigación, estructurado en dos secciones principales:

- **Sección 1:** Datos demográficos (duración estimada: 2 minutos)
    - Edad (pregunta abierta numérica)
    - Sexo biológico (pregunta de opción múltiple: Masculino/Femenino)
    - Entidad donde realiza el internado: Essalud, Minsa, Fuerzas Armadas, Clínicas privadas.
- **Sección 2:** Evaluación diagnóstica basada en imágenes clínicas (duración estimada: 20 - 30 minutos).

El cuestionario contiene 15 preguntas, cada una presentando una imagen clínica de alta resolución correspondiente a una patología dermatológica distinta. Las 15 patologías seleccionadas por prevalencia en atención primaria y relevancia clínica son:

- **Enfermedades inflamatorias:** Dermatitis atópica, Dermatitis de contacto, Psoriasis, Rosácea, Acné vulgar
- **Enfermedades infecciosas:** Impétigo, Tiña corporis, Molusco contagioso, Verruga vulgar
- **Enfermedades pigmentarias benignas:** Nevus melanocítico, Lentigo solar, Queratosis seborreica
- **Enfermedades malignas y premalignas:** Melanoma, Carcinoma basocelular, Queratosis actínica

Selección y origen de las imágenes clínicas:

Para la elaboración del cuestionario se seleccionaron 15 imágenes clínicas representativas provenientes del Atlas Dermatológico Brasileño

(https://atlasdermatologico.com.br/browse.jsf. La elección de este repositorio se fundamentó en:

- Validez clínica de las imágenes y supervisión por dermatólogos certificados
- Representatividad de diversos fototipos de piel (Fitzpatrick I-VI)
- Accesibilidad para fines académicos y educativos
- Mayor diversidad fenotípica comparada con otras bases internacionales

Se descartó inicialmente el uso de DermNet NZ debido a su mayor énfasis en dermatoscopia y menor representatividad de fototipos oscuros, así como el ISIC Archive por su enfoque predominantemente dermatoscópico.

Criterios de selección de imágenes:

- Presentaciones clínicas típicas pero no estereotipadas de cada patología
- Diversidad de fototipos cuando aplicable
- Imágenes con alta resolución y calidad fotográfica
- Ausencia de artefactos, tatuajes, cicatrices o lesiones secundarias que dificulten el reconocimiento
- Lesiones en ubicaciones anatómicas comunes para cada patología
- Sin información clínica adicional visible ni datos identificables del paciente

Características técnicas de las imágenes:

- Formato: JPEG con resolución estandarizada
- Iluminación uniforme, sin sombras ni destellos
- Presentación centrada y sin anotaciones ni marcas
- Una imagen por patología, presentadas secuencialmente

Opciones de respuesta:

- Para cada pregunta con imagen, se presentaron las 15 opciones diagnósticas en formato de lista desplegable permitiendo que el egresado seleccione una única respuesta. Las opciones se presentaron en orden alfabético para facilitar la búsqueda y evitar patrones de respuesta predecibles.

Plataforma tecnológica y herramientas de desarrollo:

El cuestionario fue desarrollado como aplicación web utilizando:

- Lenguajes de programación: HTML5, CSS3 y JavaScript (ES6+)
- Plataforma de alojamiento: Vercel (https://vercel.com), servicio gratuito de hosting para aplicaciones web estáticas
- Sistema de almacenamiento de respuestas: Google Sheets mediante integración con Google Apps Script
- Script de backend: Google Apps Script configurado para recibir peticiones POST y almacenar respuestas automáticamente en hoja de cálculo

La aplicación web presenta las siguientes características funcionales:

- Carátula inicial con título del estudio, siglas (CPDD-IC), autor y objetivo resumido
- Consentimiento informado digital con aceptación obligatoria mediante checkbox
- Navegación secuencial entre preguntas con botones "Anterior" y "Siguiente"
- Barra de progreso visual indicando avance (1/15, 2/15, etc.)
- Envío automático de respuestas al completar la última pregunta
- Bloqueo del botón "Finalizar" tras primer clic para evitar duplicación de envíos
- Tabla de resultados mostrando respuestas del participante, respuestas correctas, imágenes miniatura y precisión diagnóstica global calculada

Se adjunta enlace y formato de la encuesta en Anexo 4.

3.8.2. Validación del instrumento

El Cuestionario de Precisión Diagnóstica Dermatológica mediante Imágenes Clínicas (CPDD-IC) fue sometido a un proceso riguroso de validación en dos fases: validación de contenido por juicio de expertos y valoración psicométrica mediante estudio piloto.

Validez de contenido por juicio de expertos

Cinco dermatólogos certificados con experiencia docente en instituciones de salud de Lima (promedio de experiencia clínica: 12.4 años; rango: 8-18 años) evaluaron la relevancia, claridad, suficiencia y coherencia de cada ítem del cuestionario mediante escala Likert de 5 puntos (1 = totalmente inadecuado, 5 = totalmente adecuado).

Se calculó el Coeficiente de Validez de Contenido (CVC) para cada ítem y para el instrumento global, obteniendo valores superiores a 0.80 para todos los ítems evaluados, indicando validez de contenido adecuada (90).

Los expertos también evaluaron: (a) calidad técnica de las imágenes clínicas (resolución, iluminación, encuadre, representatividad de la patología), (b) representatividad del espectro de patologías seleccionadas respecto a prevalencia en atención primaria, (c) pertinencia de opciones diagnósticas distractoras, y (d) claridad de instrucciones. Las observaciones cualitativas recibidas fueron incorporadas en la versión final del instrumento, incluyendo ajustes en redacción de instrucciones y sustitución de dos imágenes por alternativas de mayor calidad técnica.

El cuestionario mediante escala Likert se encuentra adjunto en el Anexo 5. El procedimiento detallado de validación por expertos, incluyendo datos de filiación, matriz de calificaciones detallada de los cinco expertos, fórmulas de CVC, tablas de evaluación individual por cada experto, y análisis estadístico completo se presentan en el Anexo 6.

Análisis de confiabilidad y propiedades psicométricas mediante estudio piloto

Se realizó un estudio piloto con 30 egresados de medicina de la UNMSM pertenecientes a la misma población objetivo del estudio (miembros de la promoción 2025). Los objetivos de esta prueba preliminar fueron: (a) evaluar la confiabilidad interna del instrumento en una muestra representativa, (b) analizar la discriminación de los ítems, (c) identificar dificultades de comprensión o problemas técnicos en la plataforma, y (d) estimar el tiempo promedio de completación del cuestionario.

La confiabilidad del instrumento fue evaluada mediante coeficiente Kuder-Richardson (KR-20), apropiado para instrumentos dicotómicos (correcto/incorrecto), obteniendo KR-20 = 0.78, valor que indica consistencia interna aceptable según criterios psicométricos estándar (valores > 0.70 se consideraron adecuados para instrumentos de evaluación de competencias).

El índice de discriminación de ítems fue calculado para cada una de las 15 preguntas mediante comparación de desempeño del tercil superior versus tercil inferior de participantes. Todos los ítems mostraron índice de discriminación > 0.20, indicando que cada pregunta discrimina apropiadamente entre participantes con diferente nivel de competencia diagnóstica.

Los participantes del piloto (93%) reportaron que las instrucciones eran claras y comprensibles, y que la plataforma digital era intuitiva y de fácil navegación. No se identificaron problemas técnicos significativos durante la aplicación del piloto.

Las fórmulas estadísticas detalladas, tablas de análisis de ítems, y resultados completos del estudio piloto se presentan en Anexo 7.

Síntesis de la validación

Los resultados de las dos fases de validación (juicio de expertos y piloto psicométrico) indican que el CPDD-IC es un instrumento válido y confiable para evaluar precisión diagnóstica dermatológica mediante imágenes clínicas en egresados de medicina, cumpliendo con estándares metodológicos establecidos para instrumentos de evaluación de competencias clínicas.

Los participantes en la prueba piloto completaron el mismo consentimiento informado digital que será usado en la muestra principal, garantizando cumplimiento con normativas éticas. No se recolectó información identificable de los participantes, manteniendo el anonimato durante todo el proceso de validación.

3.9. Proceso de recolección de datos

La recolección de datos se llevó a cabo tras la obtención de la aprobación del Comité de Ética en Investigación de la UNMSM (Anexo 8). Con esta autorización, se procedió a coordinar el acceso a la cohorte de egresados 2025 y a definir los canales institucionales de comunicación.

El reclutamiento de participantes se realizó en febrero de 2026 mediante el envío de invitaciones electrónicas a los egresados, utilizando los correos institucionales registrados en la Escuela de Medicina. En estas comunicaciones se explicó el objetivo del estudio, el carácter voluntario de la participación, el tiempo estimado para completar el cuestionario (25–35 minutos) y las condiciones de confidencialidad de la información. Los egresados interesados confirmaron su participación a través de los mismos canales electrónicos.

Durante febrero de 2026 se administró el instrumento CPDD-IC en formato digital. A los participantes que confirmaron su participación se les remitió un enlace al cuestionario autoadministrado, accesible desde computadoras personales y dispositivos móviles. El cuestionario pudo completarse sin límite rígido de tiempo de sesión, de modo que cada participante respondió a su propio ritmo, en una única sesión continua.

Al cierre del período de aplicación, se procedió a la descarga de la base de datos en formato CSV. Posteriormente se realizó la verificación de integridad de los registros, así como la limpieza y codificación de las variables, obteniéndose la base de datos final utilizada para el análisis estadístico.

3.10. Análisis estadístico de los datos

Los datos recolectados se procesaron y analizaron utilizando el software SPSS versión 27.0 y Python 3.x, empleando principalmente las bibliotecas pandas y numpy para el manejo y transformación de datos, y seaborn junto con matplotlib para la generación de la matriz de confusión y su representación gráfica.

Para las características demográficas y académicas se aplicó estadística descriptiva: frecuencias absolutas y relativas para variables categóricas (sexo, sede de internado) y medidas de tendencia central y dispersión para variables numéricas (edad).

La precisión diagnóstica se cuantificó en tres niveles: global, por patología y por categoría clínica. La precisión global se calculó como la proporción de respuestas correctas sobre el total de respuestas, expresada en porcentaje. Para cada una de las 15 patologías se estimó el porcentaje de egresados que realizó el

diagnóstico correcto, y se calcularon los porcentajes de acierto promedio para cada una de las cuatro categorías clínicas (inflamatorias, infecciosas, pigmentadas benignas y premalignas/malignas).

Para el análisis de patrones de error se construyó, en Python, una matriz de confusión multiclase 15×15 (diagnóstico real vs. diagnóstico predicho) utilizando pandas para organizar los datos. A partir de esta matriz se identificaron las confusiones frecuentes, definidas como celdas fuera de la diagonal con frecuencia ≥5 egresados o proporción ≥5% respecto al total de respuestas de la patología real, y se describieron las 10 más frecuentes. Para cada par de confusión se analizó la direccionalidad (simétrica o asimétrica) y se clasificó el error como intra-grupo o inter-grupo según la categoría clínica de las patologías implicadas, estimándose además la proporción total de errores intra-grupo e inter-grupo. La matriz se representó gráficamente mediante un heatmap generado con seaborn/matplotlib, a fin de facilitar la visualización de patrones sistemáticos de error.

Para el análisis comparativo de la precisión diagnóstica se aplicó la prueba Q de Cochran para comparar la proporción de aciertos entre las 15 patologías y la prueba de Friedman para comparar la precisión entre las cuatro categorías clínicas, con análisis post-hoc de Wilcoxon y ajuste por Bonferroni cuando fue pertinente. Las comparaciones entre grupos independientes (sexo y sede de internado) se realizaron mediante pruebas t de Student o ANOVA, o sus equivalentes no paramétricos U de Mann–Whitney y Kruskal–Wallis, de acuerdo con los supuestos de normalidad y homocedasticidad. La asociación entre edad y precisión diagnóstica global se evaluó mediante correlación de Pearson o Spearman, según la distribución de las variables. Se consideró un nivel de significancia estadística de $p<0,05$ ($\alpha=0,05$).

3.11. Consideraciones éticas

El estudio contó con la aprobación previa del Comité de Ética en Investigación de la Facultad de Medicina de la UNMSM, obtenida en febrero de 2026, así como con la autorización de las autoridades académicas correspondientes. La investigación se clasificó como de riesgo mínimo, dado que la intervención se limitó a la aplicación de un cuestionario autoadministrado no

invasivo, sin manipulación de datos sensibles más allá de información demográfica básica y con participación completamente voluntaria.

La participación se realizó únicamente tras la aceptación de un consentimiento informado digital, presentado antes del acceso al cuestionario en la plataforma. Este documento incluía la explicación del objetivo del estudio, la descripción del procedimiento, la garantía de confidencialidad, el derecho a retirarse en cualquier momento sin perjuicio alguno y los datos de contacto del investigador responsable. El formulario exigía una confirmación explícita de aceptación para poder continuar.

Se garantizó el anonimato de los participantes: no se recolectaron nombres, números de documento de identidad ni otra información que permitiera la identificación directa. Los datos fueron almacenados en un servidor seguro, con acceso restringido al equipo investigador y resguardo mediante mecanismos de protección y encriptación. El modelo de consentimiento informado utilizado se presenta en el Anexo 3.

4. CAPÍTULO IV: Resultados

4.1. Características de la muestra

Se evaluó a 113 egresados de medicina humana que realizaron su internado médico en instituciones de salud de Lima. La edad promedio de los participantes fue de 25,8 años (DE: 2,3; rango: 22–33 años). La mayoría fue de sexo femenino (52,2%), con una distribución similar de varones (47,8%). En cuanto a la sede de internado, el 57,5% realizó su internado en establecimientos del MINSA, 31,9% en EsSalud y 10,6% en instituciones privadas.

Tabla 1. Características demográficas y académicas de los egresados de medicina de la UNMSM, 2025

Características	n	%
Edad (años)	25,8 ± 2,3	
Sexo		
Masculino	54	47,8
Femenino	59	52,2
Sede de internado		
MINSA	65	57,5
EsSalud	36	31,9
Clínicas privadas	12	10,6

Fuente: Elaboración propia

4.2. Precisión diagnóstica

La precisión diagnóstica global para la identificación de las 15 patologías dermatológicas fue de 70,7%, considerando el total de respuestas emitidas por los egresados. A nivel individual, la mayoría de participantes obtuvo porcentajes de aciertos superiores al 60%.

En la Tabla 2 se presenta la precisión diagnóstica para cada patología evaluada. Las patologías con mayor proporción de aciertos fueron rosácea (91,2%), acné vulgar (88,5%) y tiña corporis (84,1%). En contraste, las patologías con menor precisión fueron queratosis seborreica (24,8%), queratosis actínica (33,6%) y dermatitis atópica (51,3%).

Tabla 2. Precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 según patología dermatológica específica.

Patología	n aciertos	% aciertos
Rosácea	103	91,2
Acné vulgar	100	88,5
Tiña corporis	95	84,1
Carcinoma basocelular	95	84,1
Melanoma	94	83,2
Impétigo no bulloso	93	82,3
Molusco contagioso	91	80,5
Psoriasis	89	78,8
Nevo melanocítico común	87	77,0
Lentigo solar	82	72,6
Dermatitis de contacto alérgica	78	69,0
Verruga vulgar	67	59,3
Dermatitis atópica	58	51,3
Queratosis actínica	38	33,6
Queratosis seborreica	28	24,8

Fuente: Elaboración propia

Al agrupar las patologías según categoría clínica, se observó mayor precisión diagnóstica para patologías infecciosas (76,6%) e inflamatorias (75,8%), seguidas por premalignas/malignas (67,0%). La menor precisión se obtuvo en patologías pigmentadas benignas (58,1%).

Tabla 3. Precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 según categoría clínica de patologías dermatológicas.

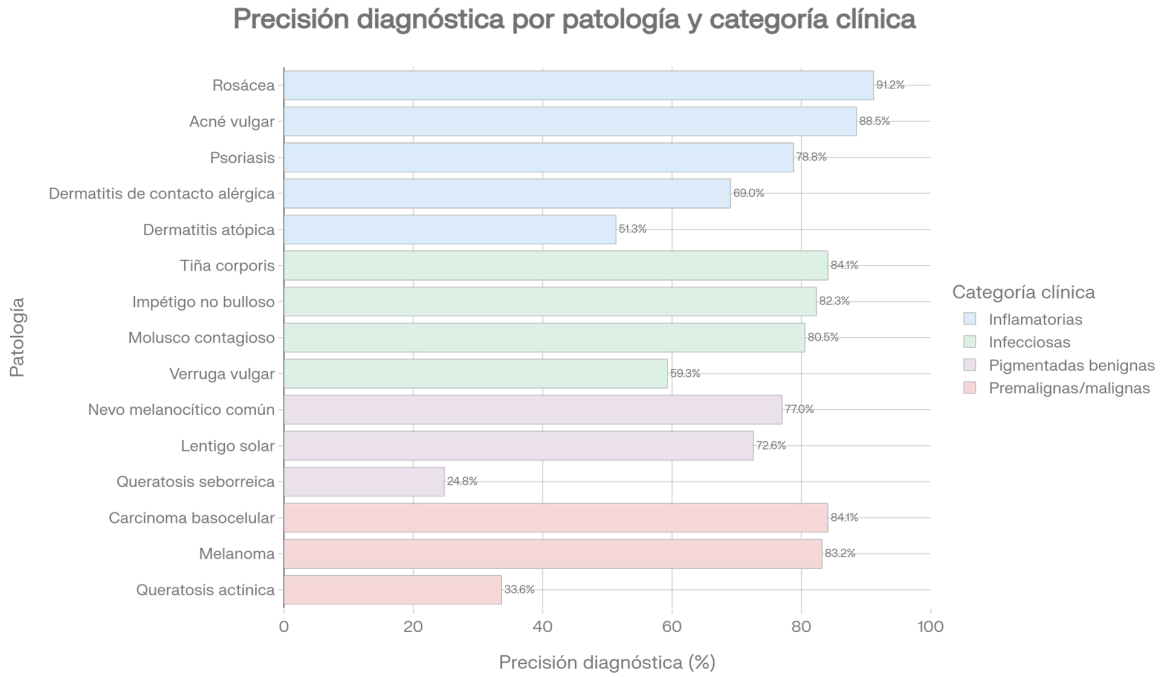
Categoría clínica	n patologías	Respuestas correctas	Respuestas totales	% precisión
Infecciosas	4	346	452	76,6
Inflamatorias	5	429	566	75,8
Premalignas/malignas	3	227	339	67,0
Pigmentadas benignas	3	197	338	58,1

Fuente: Elaboración propia

En la Figura 1 se observa la precisión diagnóstica individual de cada patología, agrupada por categoría clínica. Las patologías inflamatorias e infecciosas concentran las mayores proporciones de aciertos (por ejemplo, rosácea, acné vulgar y tiña corporis), mientras que dentro de las pigmentadas benignas y las premalignas/malignas se identifican las entidades con menor precisión, como queratosis seborreica y queratosis actínica. Esta distribución gráfica complementa lo presentado en la Tabla 2 y permite visualizar cómo los bajos porcentajes de

acierto de algunas patologías específicas contribuyen a reducir la precisión promedio de sus respectivas categorías clínicas (Tabla 3).

Figura 1. Precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 según patología y categoría clínica.

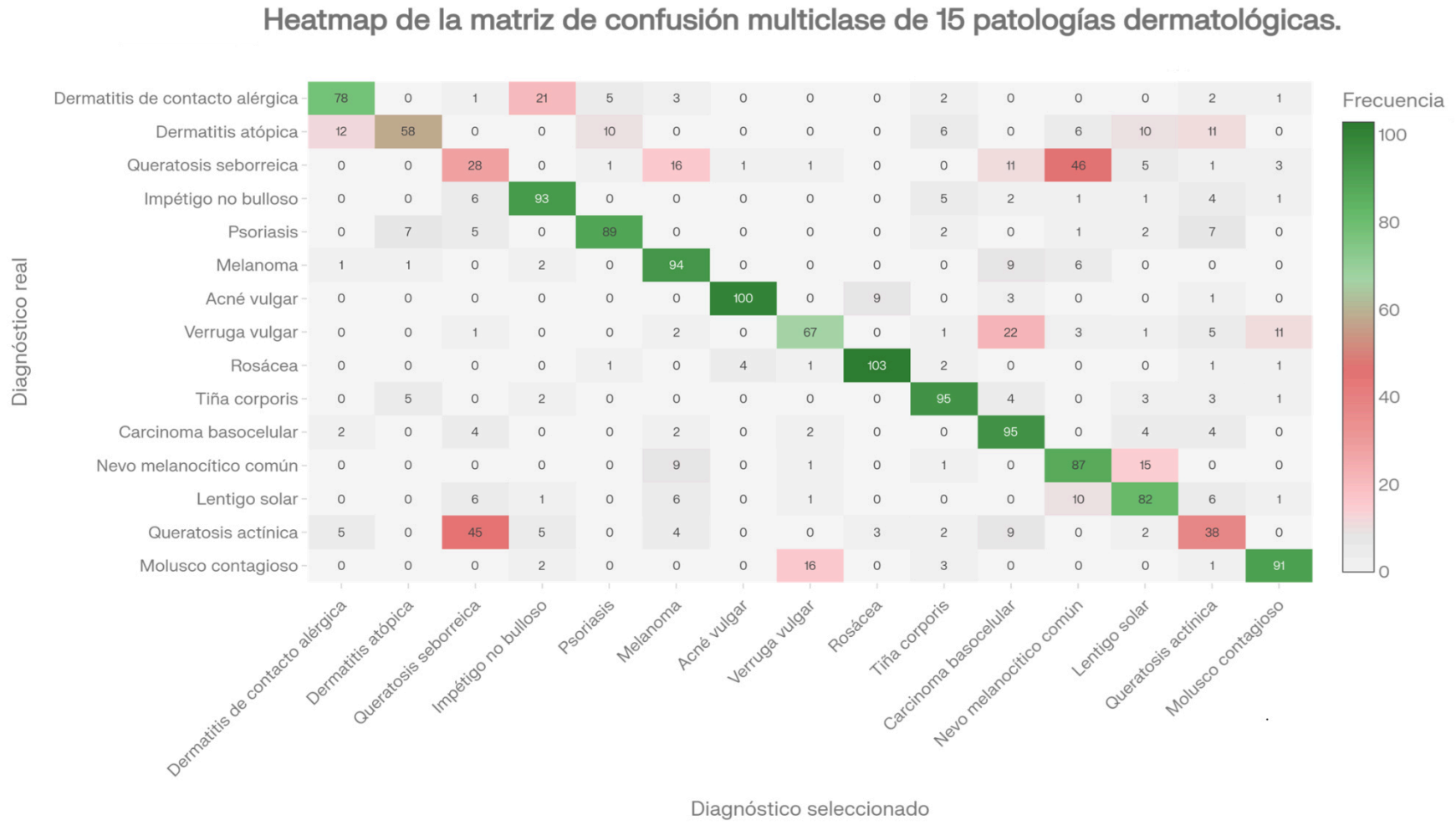


Fuente: Elaboración propia

4.3. Matriz de confusión multiclase

A partir de los diagnósticos emitidos por los egresados se construyó una matriz de confusión multiclase de dimensión 15×15, donde las filas representan las patologías verdaderas y las columnas los diagnósticos seleccionados por los participantes. La diagonal principal corresponde a los aciertos, mientras que las celdas fuera de la diagonal representan errores de clasificación específicos. Esta matriz se representa gráficamente en la Figura 2 mediante un heatmap, resaltando la diagonal principal para facilitar la visualización de los patrones sistemáticos de error.

Figura 2. Matriz de confusión multiclase de los diagnósticos dermatológicos realizados por los egresados de medicina de la UNMSM del 2025.



Fuente: Elaboración propia

Para el análisis de patrones de error se definieron como confusiones frecuentes aquellas celdas fuera de la diagonal con frecuencia ≥5 egresados o proporción ≥ 5% respecto al total de respuestas de la patología real. En la Tabla 4 se resumen las 10 confusiones diagnósticas más frecuentes. La confusión más frecuente fue queratosis seborreica diagnosticada como nevo melanocítico común (40,7%), seguida de queratosis actínica diagnosticada como queratosis seborreica (39,8%) y verruga vulgar diagnosticada como carcinoma basocelular (19,5%).

Cada confusión se clasificó como intra-grupo cuando la patología real y el diagnóstico predicho pertenecían a la misma categoría clínica, o inter-grupo cuando correspondían a categorías diferentes. En el conjunto de errores de clasificación de la matriz de confusión, el 46–47% correspondió a errores intra-grupo y aproximadamente 53–54% a errores inter-grupo, mostrando un ligero predominio de confusiones entre categorías clínicas distintas. Dentro de las 10 confusiones más frecuentes (Tabla 4), cinco fueron intra-grupo y cinco inter-grupo, lo que evidencia tanto dificultad para discriminar entidades morfológicamente similares dentro de una misma categoría como errores más groseros entre grupos clínicamente distintos.

Tabla 4. Patrones de error: principales confusiones diagnósticas de los egresados de medicina de la UNMSM del 2025.

Patología real	Diagnóstico predicho	n errores	% sobre fila	Tipo de error
Queratosis seborreica	Nevo melanocítico común	46	40,7	Intra-grupo
Queratosis actínica	Queratosis seborreica	45	39,8	Inter-grupo
Verruga vulgar	Carcinoma basocelular	22	19,5	Inter-grupo
Dermatitis de contacto alérgica	Impétigo no bulloso	21	18,6	Inter-grupo
Queratosis seborreica	Melanoma	16	14,2	Inter-grupo
Molusco contagioso	Verruga vulgar	16	14,2	Intra-grupo
Nevo melanocítico común	Lentigo solar	15	13,3	Intra-grupo
Dermatitis atópica	Dermatitis de contacto alérgica	12	10,6	Intra-grupo
Verruga vulgar	Molusco contagioso	11	9,7	Intra-grupo
Dermatitis atópica	Queratosis actínica	11	9,7	Inter-grupo

Fuente: Elaboración propia

Para cada uno de los pares de patologías incluidos en la Tabla 4 se evaluó la direccionalidad de la confusión, comparando la frecuencia de A confundida con B (A→B) frente a B confundida con A (B→A). En la mayoría de los pares, la confusión fue asimétrica, es decir, una dirección predominó claramente sobre la otra. Por ejemplo, queratosis seborreica fue confundida con nevo melanocítico común con mucha mayor frecuencia que la inversa, y queratosis actínica fue confundida con queratosis seborreica más frecuentemente que a la inversa. Esto sugiere que, en términos pedagógicos, ciertas patologías funcionan como "diagnóstico refugio" o categoría más accesible para el egresado cuando no reconoce adecuadamente la lesión real. En la Tabla 5 se ilustra la direccionalidad de algunas de las confusiones más frecuentes.

Tabla 5. Patrones de error: direccionalidad de la confusión diagnóstica de los egresados de medicina de la UNMSM del 2025.

Par de patologías	Dirección de confusión	n	Interpretación de la direccionalidad
Queratosis seborreica (QS) y Nevo melanocítico común (N)	QS → N N → QS	46 0	Predomina la confusión de queratosis seborreica hacia nevo melanocítico común (asimétrica)
Queratosis actínica (QA) y Queratosis seborreica (QS)	QA → QS QS → QA	45 0	Predomina la confusión de queratosis actínica hacia queratosis seborreica (asimétrica)
Verruga vulgar (VV) y Molusco contagioso (MC)	VV → MC MC → VV	11 16	La confusión se concentra principalmente en la dirección verruga vulgar hacia molusco contagioso (asimétrica)

Fuente: Elaboración propia

4.4. Análisis comparativo de la precisión diagnóstica

Comparación entre patologías y categorías clínicas

Al aplicar la prueba Q de Cochran para comparar la proporción de aciertos entre las 15 patologías, se encontraron diferencias estadísticamente significativas

en la precisión diagnóstica global (Q=343,1; p<0,001). Este resultado indica que las proporciones de aciertos no son homogéneas entre las patologías, es decir, algunas resultan considerablemente más fáciles o más difíciles de identificar que otras. En concordancia con la Tabla 2, las mayores proporciones de aciertos se observaron en rosácea, acné vulgar y tiña corporis, mientras que las menores corresponden a queratosis seborreica, queratosis actínica y dermatitis atópica.

De forma similar, al agrupar las patologías en cuatro categorías clínicas y comparar la precisión promedio mediante la prueba de Friedman, también se evidenciaron diferencias estadísticamente significativas (χ^2=60,9; p<0,001). Este resultado indica que la competencia diagnóstica de los egresados varía según el tipo de entidad dermatológica (inflamatoria, infecciosa, pigmentada benigna o premaligna/maligna), y no es homogénea entre categorías.

En el análisis post-hoc para las categorías clínicas (Tabla 6), utilizando la prueba de rangos con signo de Wilcoxon con ajuste de Bonferroni (α≈0,008), no se encontraron diferencias significativas entre las categorías inflamatorias e infecciosas (p=0,384), lo que sugiere un desempeño similar entre ambas. Sin embargo, inflamatorias e infecciosas mostraron mayor precisión que las pigmentadas benignas y las premalignas/malignas (todas las comparaciones p<0,001), lo que evidencia un menor rendimiento en el reconocimiento de lesiones pigmentadas. Además, la categoría de pigmentadas benignas presentó precisión significativamente menor que la de lesiones premalignas/malignas (p<0,001), señalando un área de mayor dificultad diagnóstica.

Este resultado indica que la competencia diagnóstica de los egresados varía según el tipo de entidad dermatológica (inflamatoria, infecciosa, pigmentada benigna o premaligna/maligna), y no es homogénea entre categorías. Por lo tanto, al obtener un valor p < 0,001, se confirma estadísticamente la hipótesis de investigación planteada, demostrando que existen diferencias significativas en la precisión diagnóstica entre las distintas categorías clínicas evaluadas.

Tabla 6. Comparaciones post-hoc de la precisión diagnóstica de los egresados de medicina de la UNMSM del 2025 entre categorías clínicas dermatológicas.

Comparación entre categorías clínicas	p-valor (Wilcoxon)[1]
Inflamatorias vs Infecciosas	0,384
Inflamatorias vs Pigmentadas benignas	<0,001
Inflamatorias vs Premalignas/malignas	<0,001
Infecciosas vs Pigmentadas benignas	<0,001
Infecciosas vs Premalignas/malignas	<0,001
Pigmentadas benignas vs Premalignas/malignas	<0,001

[1] Prueba de rangos con signo de Wilcoxon, con ajuste de Bonferroni para 6 comparaciones (α≈0,008).
Fuente: Elaboración propia

Precisión diagnóstica según sexo

En la Tabla 7 se presenta la precisión diagnóstica global según sexo. Las egresadas alcanzaron una precisión media de 68,8%, mientras que los egresados obtuvieron 72,7%. Aunque el promedio fue ligeramente mayor en varones, la diferencia no fue estadísticamente significativa (U de Mann–Whitney=1333,5; p=0,134). Estos resultados indican que, en esta cohorte, el rendimiento diagnóstico global en dermatología es comparable entre hombres y mujeres.

Tabla 7. Precisión diagnóstica global de los egresados de medicina de la UNMSM del 2025 según sexo.

Sexo	n	Precisión global media (%)	p-valor[1]
Femenino	59	68,8	
Masculino	54	72,7	0,134

[1] Prueba U de Mann−Whitney.
Fuente: Elaboración propia

Precisión diagnóstica según sede de internado

La Tabla 8 muestra la precisión diagnóstica global según sede de internado. Los egresados procedentes de instituciones privadas presentaron la mayor precisión (80,0%), seguidos por EsSalud (73,2%) y MINSA (67,6%). No obstante, estas diferencias no alcanzaron significancia estadística en la prueba de Kruskal–Wallis (H=4,95; p=0,084). Este patrón sugiere una posible tendencia a mejor rendimiento

en algunos contextos formativos, aunque los resultados no permiten afirmar diferencias concluyentes entre sedes.

Tabla 8. Precisión diagnóstica global de los egresados de medicina de la UNMSM del 2025 según sede de internado.

Sede de internado	n	Precisión global media (%)	p-valor[1]
MINSA	65	67,6	
EsSalud	36	73,2	
Privadas	12	80,0	0,084

[1] Prueba de Kruskal–Wallis.
Fuente: Elaboración propia

Correlación entre edad y precisión diagnóstica

Finalmente, se evaluó la relación entre edad y precisión diagnóstica global. El coeficiente de Spearman fue bajo y no significativo (rho = 0,092; p = 0,332), lo que indica ausencia de asociación relevante entre la edad del egresado y su rendimiento diagnóstico en dermatología dentro del rango etario estudiado.

Al analizar de forma conjunta los resultados obtenidos según sexo (p = 0,134), sede de internado (p = 0,084) y edad (p = 0,332), se evidencia que ningún valor p fue menor al nivel de significancia del 5% (0,05). En consecuencia, no se halló evidencia estadística suficiente para respaldar la hipótesis de investigación, concluyendo que no existen diferencias significativas en la precisión diagnóstica según las características sociodemográficas y académicas de los egresados evaluados.

5. CAPÍTULO V: Discusión

La precisión diagnóstica global cercana al 70% observada en este estudio indica que los egresados de medicina de la UNMSM son capaces de identificar correctamente la mayoría de las patologías dermatológicas incluidas en el set de 15 imágenes clínicas, pero con un rendimiento heterogéneo según la entidad evaluada. Este hallazgo tiene especial relevancia en el contexto peruano, donde no existían mediciones objetivas de competencia diagnóstica visual en dermatología en estudiantes o egresados, a pesar de los reportes de diagnóstico tardío en patología cutánea y de las dificultades estructurales del sistema de salud para garantizar atención oportuna (31). En ese escenario, los resultados ofrecen por primera vez una línea de base cuantitativa sobre lo que los egresados realmente pueden hacer frente a imágenes clínicas concretas, y no solo sobre lo que perciben saber.

Al analizar la precisión por patología, se observa que el desempeño no es uniforme. Patologías como rosácea, acné vulgar, tiña corporis, carcinoma basocelular y melanoma alcanzan proporciones de acierto elevadas, mientras que queratosis seborreica, queratosis actínica, dermatitis atópica y verruga vulgar presentan porcentajes considerablemente menores. La caída drástica en la precisión de estas entidades específicas encuentra su justificación en la propia fenomenología del diagnóstico dermatológico. La dermatitis atópica, por ejemplo, es un diagnóstico eminentemente sindrómico que depende de la cronicidad, el prurito severo y la topografía generalizada de las lesiones, elementos invisibles en una imagen estática y focalizada. Por su parte, la queratosis actínica es una entidad donde la palpación (el característico tacto 'en papel de lija') es frecuentemente más diagnóstica que la inspección visual pura. El bajo rendimiento en estas patologías evidencia que, cuando el médico general es privado del componente táctil y anamnésico, su capacidad para discriminar visualmente estas lesiones disminuye de forma drástica.

La prueba Q de Cochran mostró diferencias estadísticamente significativas entre las 15 patologías, lo que confirma que no todas representan el mismo nivel de dificultad diagnóstica. Desde una perspectiva formativa, esto sugiere que los egresados han internalizado con relativa solidez los patrones clínicos de ciertas entidades frecuentes y “prototípicas”, mientras que mantienen debilidades en

cuadros donde el diagnóstico diferencial visual exige distinguir matices más sutiles. Resultados de intervenciones educativas breves apoyan esta interpretación: en el estudio cuasiexperimental de Ricco et al. se demostró que una sesión estructurada de 60 minutos fue suficiente para mejorar de manera significativa la descripción morfológica, la precisión diagnóstica y la confianza ante imágenes de patologías comunes (26). Los datos del presente trabajo, sin embargo, indican que esa consolidación aún no ocurre de manera equilibrada en todas las áreas de la dermatología durante el pregrado.

La comparación por categorías clínicas refuerza la existencia de focos de debilidad específicos. La precisión en patologías inflamatorias e infecciosas fue alta y similar, mientras que el rendimiento disminuyó de forma notable en pigmentadas benignas y en lesiones premalignas/malignas. La prueba de Friedman evidenció diferencias significativas entre las cuatro categorías, y los análisis post-hoc mostraron que las inflamatorias e infecciosas superan de forma consistente a las categorías pigmentadas. Esta asimetría puede interpretarse, al menos en parte, como resultado de la exposición clínica: las infecciones y dermatosis inflamatorias son muy frecuentes en atención primaria, lo que favorece su reconocimiento, mientras que el entrenamiento sistemático en evaluación de patrones pigmentados (bordes, color, asimetría, distribución) suele ser más limitado en el pregrado.

Adicionalmente, estos resultados revelan una especie de 'paradoja educativa' en la enseñanza de la dermatología oncológica. Mientras que el melanoma alcanzó una precisión notablemente alta (83,2%), las lesiones benignas con las que suele hacer diagnóstico diferencial, como la queratosis seborreica, obtuvieron los puntajes más bajos. Esto sugiere que el currículo médico enfatiza exitosamente las señales de alarma (red flags) y la regla ABCDE por el justificado temor a omitir un cáncer letal, pero no entrena con el mismo rigor el reconocimiento de los patrones benignos comunes. Como consecuencia, el egresado sabe identificar la malignidad franca, pero carece de las herramientas visuales para confirmar la benignidad, lo que en la práctica clínica se traduciría en sobrediagnóstico, ansiedad en el paciente y derivaciones innecesarias al nivel especializado.

El consenso regional de Sánchez et al. describe precisamente barreras educativas y estructurales en dermatología en América Latina, incluyendo

conocimiento inadecuado de enfermedades dermatológicas comunes entre médicos de atención primaria, déficits en la capacitación clínica y distribución desigual de recursos y especialistas (28). El presente estudio no demuestra directamente estos mecanismos, pero los déficits observados en lesiones pigmentadas son coherentes con un contexto donde la formación dermatológica no se ha priorizado ni estructurado de manera uniforme, y donde los contenidos más complejos probablemente reciben menos énfasis práctico.

El análisis de la matriz de confusión complementa estos hallazgos al mostrar patrones concretos de error. Las confusiones más frecuentes (queratosis seborreica interpretada como nevo melanocítico, queratosis actínica como queratosis seborreica y verruga vulgar como carcinoma basocelular) no son aleatorias, sino que siguen direcciones bien definidas. La asimetría de estas confusiones sugiere que, frente a imágenes ambiguas, los egresados tienden a reclasificar las lesiones hacia diagnósticos que perciben como más familiares o que funcionan como categorías amplias, como nevos melanocíticos o queratosis seborreica.

En algunos casos, como la alarmante confusión de verruga vulgar hacia carcinoma basocelular, la dirección del error no solo refleja una deficiencia morfológica, sino que ilustra un claro sesgo cognitivo de aversión al riesgo (o actitud precautoria). Ante la incertidumbre visual, el médico recién egresado tiende a anclar su decisión en el diagnóstico más grave (el peor escenario posible) para proteger al paciente. Si bien esto es una red de seguridad desde el punto de vista médico-legal, revela que los esquemas mentales de enfermedad no están lo suficientemente maduros para diferenciar un tumor maligno de una hiperplasia viral común. El hecho de que las confusiones se concentren en determinados pares asimétricos indica que, ante la falta de heurísticas visuales eficientes, la incertidumbre se canaliza sistemáticamente hacia estos “diagnósticos refugio”, lo cual acarrea potenciales consecuencias clínicas y derivaciones innecesarias en el nivel primario de atención.

La proporción de errores intra-grupo e inter-grupo agrega otra dimensión al análisis. Una fracción importante de las equivocaciones se produce entre patologías de la misma categoría clínica, lo cual es esperable en egresados sin entrenamiento especializado y apunta a la necesidad de mejorar la discriminación visual fina dentro de cada grupo. No obstante, los errores inter-grupo (por ejemplo, confundir una

lesión verrugosa con un carcinoma basocelular, o no ubicar correctamente una patología en el grupo inflamatorio, infeccioso o tumoral) resultan más preocupantes, pues reflejan fallas en la estructura conceptual del razonamiento diagnóstico, no solo en la apreciación de detalles morfológicos. Estas discrepancias son coherentes con la brecha entre médicos generales y dermatólogos descrita en un metaanálisis de teledermatología, donde el acuerdo diagnóstico entre generalistas y especialistas fue considerablemente menor que el acuerdo entre dos dermatólogos evaluando las mismas imágenes (29). El presente estudio sugiere que una parte de esa brecha puede originarse ya en el pregrado, en la forma en que se enseñan y practican las categorías diagnósticas mayores.

Los análisis según sexo, sede de internado y edad no mostraron diferencias estadísticamente significativas en la precisión diagnóstica global. La ligera tendencia a mejores resultados en egresados varones o procedentes de clínicas privadas no puede interpretarse como evidencia robusta, tanto por la falta de significancia estadística como por el tamaño limitado de algunos subgrupos. Aun así, el patrón sugiere hipótesis plausibles. Por ejemplo, es razonable plantear que ciertos escenarios formativos (mayor exposición a dermatología, mejor supervisión, casuística diferente) podrían favorecer un rendimiento algo mayor, en línea con lo que Sánchez et al. describen sobre la concentración de recursos y especialistas en determinadas áreas, y las dificultades de acceso a dermatólogos en otras (28). Confirmar o refutar estas hipótesis requerirá estudios multicéntricos y muestras mayores; en el contexto actual, la lectura más prudente es considerar estos resultados como señales exploratorias, no como conclusiones firmes sobre inequidad formativa. La ausencia de correlación entre edad y precisión, por su parte, apunta a que, dentro de una cohorte de egresados recientes, la variabilidad en desempeño se explica más por las experiencias formativas concretas que por la edad cronológica, lo que coincide con hallazgos de intervenciones breves donde la mejora en competencias depende sobre todo del diseño instruccional y no tanto del año académico (26).

La interpretación de estos hallazgos debe realizarse reconociendo las limitaciones inherentes al diseño metodológico. En primer lugar, la naturaleza transversal del estudio y la participación voluntaria introducen un posible sesgo de autoselección, limitando la generalización absoluta de las inferencias a toda la

cohorte. Asimismo, al circunscribirse la muestra a egresados de la UNMSM, la validez externa se restringe a contextos formativos similares. No obstante, frente a la ausencia histórica de mediciones objetivas en el pregrado peruano, esta aproximación metodológica fue necesaria para establecer una primera línea de base cuantitativa rigurosa, sirviendo como un estudio piloto fundamental que justifica y viabiliza la futura ejecución de investigaciones multicéntricas a nivel nacional.

En segundo lugar, el uso de un instrumento basado en fotografías bidimensionales estáticas y sin acceso a la historia clínica omite elementos cruciales de la práctica real, como la palpación de las lesiones, la evolución temporal y el contexto anamnésico. A esto se suma la evaluación de un número acotado de 15 patologías, lo que no abarca la variabilidad fenotípica completa de la dermatología. Para mitigar esta restricción, se garantizó que el cuestionario visual fuera sometido a un estricto proceso de validación psicométrica por juicio de expertos y análisis de confiabilidad en un estudio piloto. Esta decisión permitió aislar y medir de manera estandarizada y objetiva la "competencia visual pura" (el reconocimiento de patrones morfológicos), eliminando variables de confusión y demostrando que, incluso en un escenario simplificado, persisten vacíos significativos en la discriminación de lesiones pigmentadas que requieren atención educativa prioritaria.

Asimismo, debe considerarse que el carácter autoadministrado de la evaluación y la ausencia de un límite de tiempo estricto para su resolución pudieron haber facilitado la consulta de fuentes externas o búsquedas en internet por parte de los participantes, lo que constituye un factor que podría sobreestimar la precisión diagnóstica real.

En síntesis, los datos obtenidos muestran que los egresados de la UNMSM alcanzan un nivel global de precisión diagnóstica aceptable en dermatología mediante imágenes clínicas, pero con un perfil claramente desigual: sólido en patologías inflamatorias e infecciosas comunes y frágil en lesiones pigmentadas benignas y premalignas. Esta debilidad se refleja tanto en las proporciones de acierto como en los patrones de confusión de la matriz, y adquiere especial relevancia en un contexto donde se han documentado diagnósticos tardíos de cáncer de piel (31). Interpretados a la luz de las barreras educativas y estructurales descritas a nivel regional (28), estos hallazgos refuerzan la necesidad de replantear

la enseñanza de dermatología en el pregrado, privilegiando entrenamientos dirigidos en diagnóstico diferencial visual de lesiones pigmentadas, el uso de imágenes clínicas estandarizadas con retroalimentación estructurada y la incorporación de instrumentos de evaluación que reflejen la complejidad real de la práctica clínica. Más que cuestionar globalmente la formación recibida, los resultados ofrecen un mapa concreto de fortalezas y debilidades sobre el cual diseñar intervenciones educativas más focalizadas y, sobre todo, más alineadas con las necesidades dermatológicas de la población que estos egresados deberán atender.

6. CAPÍTULO VI: Conclusiones y recomendaciones

Conclusiones:

- El nivel de precisión diagnóstica global en la identificación de patologías dermatológicas mediante imágenes clínicas en los egresados de medicina de la UNMSM del 2025 fue cercano al 70%. Esto indica un nivel de desempeño visual aceptable, pero perfectible para la práctica en el primer nivel de atención.

- La precisión diagnóstica específica resultó marcadamente heterogénea entre las patologías evaluadas. Entidades como rosácea, acné vulgar, tiña corporis, carcinoma basocelular y melanoma presentaron altos porcentajes de acierto, mientras que la queratosis seborreica, queratosis actínica, dermatitis atópica y verruga vulgar mostraron desempeños sustancialmente menores.

- Al comparar por categorías clínicas, se concluye que los egresados identifican con mayor precisión las patologías inflamatorias e infecciosas frente a las pigmentadas benignas y premalignas/malignas, evidenciándose un foco de debilidad formativo específico en el espectro de lesiones pigmentadas.

- El análisis de la matriz de confusión demostró que los patrones de error diagnóstico son sistemáticos y direccionales. Existen confusiones asimétricas frecuentes (especialmente entre lesiones pigmentadas y entre verrugosas/tumorales) que sugieren el uso recurrente de diagnósticos “refugio” ante la incertidumbre visual y la aversión al riesgo.

- No se evidenciaron diferencias estadísticamente significativas en la precisión diagnóstica según el sexo, la edad y la sede de internado de los participantes. Esto sugiere que las variaciones en el desempeño se explican fundamentalmente por el modelo formativo transversal de la facultad y no por características sociodemográficas.

Recomendaciones

- Integrar de forma sistemática el uso de bancos de imágenes clínicas en las evaluaciones teóricas y prácticas de los cursos de dermatología durante el pregrado, asegurando que la competencia diagnóstica visual se califique con el mismo rigor que el conocimiento teórico tradicional.

- Fomentar la inclusión y el entrenamiento con viñetas clínicas basadas en imágenes fotográficas estandarizadas dentro de los programas de preparación y simulacros para el Examen Nacional de Medicina (ENAM) y el Examen Nacional de Residentado Médico, acercando la evaluación a los desafíos reales de la inspección visual que enfrentará el médico general.

- Rediseñar los sílabos y las actividades prácticas del curso de dermatología para dedicar mayor tiempo de entrenamiento clínico y visual a las patologías específicas que demostraron los desempeños más bajos (queratosis seborreica, queratosis actínica, dermatitis atópica y verruga vulgar), garantizando la exposición del alumno a una amplia variedad de presentaciones morfológicas de estas entidades.

- Reforzar de manera prioritaria la enseñanza diferencial del espectro de lesiones pigmentadas (benignas, premalignas y malignas) mediante módulos específicos. El objetivo debe ser dotar al egresado de heurísticas visuales sólidas para confirmar la benignidad, evitando que su formación dependa exclusivamente de la identificación de signos de alarma oncológicos.

- Diseñar e implementar intervenciones educativas breves y repetidas enfocadas exclusivamente en los pares de patologías más frecuentemente confundidos. Se sugiere a los docentes de pregrado utilizar la matriz de confusión de este estudio como mapa de ruta para corregir tempranamente los sesgos cognitivos y evitar el uso de “diagnósticos refugio”.

- Promover la réplica de este diseño de evaluación estandarizada mediante estudios multicéntricos en otras facultades de medicina del país y en cohortes futuras. Esto permitirá contrastar los presentes hallazgos con distintos contextos curriculares y generar evidencia nacional que impulse reformas basadas en datos reales sobre el desempeño diagnóstico.

REFERENCIAS BIBLIOGRÁFICAS

1. Yousef H, Alhajj M, Fakoya AO, Sharma S. Anatomy, Skin (Integument), Epidermis. En: StatPearls. [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 [citado 24 de octubre de 2025]. Recuperado: http://www.ncbi.nlm.nih.gov/books/NBK470464/

2. Kim JY, Dao H. Physiology, Integument. En: StatPearls. [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 [citado 25 de octubre de 2025]. Recuperado: http://www.ncbi.nlm.nih.gov/books/NBK554386/

3. Alotaibi HM, Alruwaili ZM, Dilli AA, Altaleb AA, Asiri MM, Alwadani OJ, et al. Assessment of Primary Care Physicians' Expertise of Common Dermatological Conditions in the Jouf Region, Saudi Arabia: A Mixed Methods Study. Healthcare. 2023;11(12):1705. DOI: 10.3390/healthcare11121705

4. Richard MA, Paul C, Nijsten T, Gisondi P, Salavastru C, Taieb C, et al. The journey of patients with skin diseases from the first consultation to the diagnosis in a representative sample of the European general population from the EADV burden of skin diseases study. J Eur Acad Dermatol Venereol. 2023;37(S7):17-24. DOI: 10.1111/jdv.18916

5. Ferrari AJ, Santomauro DF, Aali A, Abate YH, Abbafati C, Abbastabar H, et al. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. 2024;403(10440):2133-61. DOI: 10.1016/S0140-6736(24)00757-8

6. Li D, Fan S, Zhao H, Song J, Li W, Xu X. Global, regional and national burden of skin and subcutaneous diseases: a systematic analysis of the Global Burden of Disease Study 2021. Int Health. 2025;ihaf070. DOI: 10.1093/inthealth/ihaf070

7. Li D, Fan S, Zhao H, Song J, Guo L, Li W, et al. Worldwide trends and future projections of fungal skin disease burden: a comprehensive analysis from the Global Burden of Diseases study 2021. Front Public Health. 2025;13:1580221. DOI: 10.3389/fpubh.2025.1580221

8. Li D, Chen M, Li W, Xu X, Li Q. Global burden of viral skin diseases from 1990 to 2021: a systematic analysis for the global burden of disease study 2021. Front Public Health. 2025;13:1464372. DOI: 10.3389/fpubh.2025.1464372

9. Milani-Nejad N, Zhang M, Kaffenberger BH. Association of Dermatology Consultations With Patient Care Outcomes in Hospitalized Patients With Inflammatory Skin Diseases. JAMA Dermatol. 2017;153(6):523. DOI: 10.1001/jamadermatol.2016.6130

10. Osman W, Ibrahim M, Yaqub S, Ibrahim M, Abdalla H, Aljaili G, et al. Enhancing dermatology education: a short-course intervention for final-year medical students. BMC Med Educ. 2025;25(1):1053. DOI: 10.1186/s12909-025-07649-y

11. Hamid I, Vandormael I, Atwan A, Mills C. Primary Care Skin Lesion Referrals Assessed via Teledermatology: A Retrospective Analysis. Cureus. 2025; DOI: 10.7759/cureus.90849

12. Karthikeyan K, Prarthana M. Competency based medical education in dermatology: Undergraduate curriculum - a cauldron of commendations, contradictions and controversies. Indian J Dermatol Venereol Leprol. 2022;88(3):282-5. DOI: 10.25259/IJDVL_216_2021

13. Trejić S, Peters HJG, Lubeek SFK, Van De Laar FA. Diagnostic Accuracy of Skin Cancer by Family Physicians. J Am Board Fam Med. 2021;34(5):984-90. DOI: 10.3122/jabfm.2021.05.210076

14. Winkler JK, Blum A, Kommoss K, Enk A, Toberer F, Rosenberger A, et al. Assessment of Diagnostic Performance of Dermatologists Cooperating With a Convolutional Neural Network in a Prospective Clinical Study: Human With Machine. JAMA Dermatol. 2023;159(6):621-7. DOI: 10.1001/jamadermatol.2023.0905

15. Mangion SE, Phan TA, Zagarella S, Cook D, Ganda K, Maibach HI. Medical school dermatology education: a scoping review. Clin Exp Dermatol. 2023;48(6):648-59. DOI: 10.1093/ced/llad052

16. Murias Quintana E, Costilla García SM, Curbelo García JJ, Calvo Blanco J, Vega Villar J, Baladrón Romero J. Analysis and technical aspects of image-related questions in the exam for Medical Internal Residence in Spain, 2022. Radiol Engl Ed. 2024;66(5):489-96. DOI: 10.1016/j.rxeng.2023.06.010

17. Thakkar S, Chavda P, Vahora R, Patel R. Image-based assessment in undergraduate dermatology curriculum: A step toward competency-based medical education. Med J Armed Forces India. 2021;77:S134-9. DOI: 10.1016/j.mjafi.2020.11.005

18. Assunção DM, Moura AS, Ribeiro LMC. VISUAL ANALOGIES IN DERMATOLOGY: IMPACT ON THE DIAGNOSTIC ACUMEN OF UNDERGRADUATE MEDICAL STUDENTS. Eur J Educ Stud. 2022;9(7). DOI: 10.46827/ejes.v9i7.4351

19. Flores-Cohaila JA, Miranda-Chavez B, Copaja-Corzo C, Benavente-Chalco XC, Rios-García W, Moreno-Ccama VP, et al. Peruvian medical residency selection: a portrayal of scores, distribution, and predictors of 28,872 applicants between 2019 and 2023. BMC Med Educ. 2025;25(1):51. DOI: 10.1186/s12909-025-06665-2

20. Garrison ZR, Hall CM, Fey RM, Clister T, Khan N, Nichols R, et al. Advances in Early Detection of Melanoma and the Future of At-Home Testing. Life. 2023;13(4):974. DOI: 10.3390/life13040974

21. Perez M, Williams N, Avila A, Bakos R, Bittencourt F, Carlos-Ortega B, et al. Dermoscopy in Selected Latin American Countries: A Preliminary Look into Current Trends and Future Opportunities Among Dermatology Residency Programs. Dermatol Pract Concept. 2023;e2023093. DOI: 10.5826/dpc.1302a93

22. Maul LV, Jamiolkowski D, Lapides RA, Mueller AM, Hauschild A, Garbe C, et al. Health Economic Consequences Associated With COVID-19–Related Delay in Melanoma Diagnosis in Europe. JAMA Netw Open. 2024;7(2):e2356479. DOI: 10.1001/jamanetworkopen.2023.56479

23. Jones K, Lennon E, McCathie K, Millar A, Isles C, McFadyen A, et al. Teledermatology to reduce face-to-face appointments in general practice during the COVID-19 pandemic: a quality improvement project. BMJ Open Qual. 2022;11(2):e001789. DOI: 10.1136/bmjoq-2021-001789

24. Krefting F, Moelleken M, Hölsken S, Placke J-M, Eisenburger RT, Albrecht LJ, et al. Comparison of visual diagnostic accuracy of dermatologists practicing in Germany in patients with light skin and skin of color. Sci Rep. 2024;14(1):8740. DOI: 10.1038/s41598-024-59426-4

25. Choy SP, Kim BJ, Paolino A, Tan WR, Lim SML, Seo J, et al. Systematic review of deep learning image analyses for the diagnosis and monitoring of skin disease. Npj Digit Med. 2023;6(1):180. DOI: 10.1038/s41746-023-00914-8

26. Ricco C, Rao BK, Pappert AS, Coppola KM. Brief Teaching Intervention Improves Medical Students' Dermatology Diagnostic Skills and Comfort in Performing Dermatology Exams. Healthc Basel Switz. 2024;12(14):1453. DOI: 10.3390/healthcare12141453

27. Rojas C, Valenzuela F, Folch H. Evaluation and comparison of the dermatology program for medical students at the University of Chile with other national and foreign universities. J Educ Health Promot. 2021;10(1). DOI: 10.4103/jehp.jehp_777_20

28. Sánchez J, Ale I-S, Angles MV, Fogelbach GG, Jansen AM, Takaoka R, et al. Healthcare Disparities in Atopic Dermatitis in Latin America: A Narrative Review. Dermatol Ther. 2023;13(2):399-416. DOI: 10.1007/s13555-022-00875-y

29. Bourkas AN, Barone N, Bourkas MEC, Mannarino M, Fraser RDJ, Lorincz A, et al. Diagnostic reliability in teledermatology: a systematic review and a meta-analysis. BMJ Open. 2023;13(8):e068207. DOI: 10.1136/bmjopen-2022-068207

30. Gutierrez EL, Galarza C, Ramos W, Tello M, Jiménez G, Ronceros G, et al. Skin diseases in the Peruvian Amazonia. Int J Dermatol. 2010;49(7):794-800. DOI: 10.1111/j.1365-4632.2010.04473.x

31. Ziegler-Rodriguez G, De La Cruz-Ku G, Desai A, Delgado LP, Maldonado S, Guart JA, et al. Carrasco AGM, editor. Melanoma in Peru: 1000 patients and 10 years of experience. PLOS One. 2025;20(10):e0323636. DOI: 10.1371/journal.pone.0323636

32. He Z, Marrone G, Ou A, Liu H, Ma L, Huang Y, et al. Factors affecting health-related quality of life in patients with skin disease: cross-sectional results from 8,789 patients with 16 skin diseases. Health Qual Life Outcomes. 2020;18(1):298. DOI: 10.1186/s12955-020-01542-6

33. Sanclemente G, Burgos C, Nova J, Hernández F, González C, Reyes MI, et al. The impact of skin diseases on quality of life: A multicenter study. Actas Dermo-Sifiliográficas. 2017;108(3):244-52. DOI: 10.1016/j.ad.2016.11.008

34. Picardi A, Lega I, Tarolla E. Suicide risk in skin disorders. Clin Dermatol. 2013;31(1):47-56. DOI: 10.1016/j.clindermatol.2011.11.006

35. Ghosh S, Behere R, Sharma P, Sreejayan K. Psychiatric evaluation in dermatology: An overview. Indian J Dermatol. 2013;58(1):39. DOI: 10.4103/0019-5154.105286

36. Lim HW, Collins SAB, Resneck JS, Bolognia JL, Hodge JA, Rohrer TA, et al. The burden of skin disease in the United States. J Am Acad Dermatol. 2017;76(5):958-972.e2. DOI: 10.1016/j.jaad.2016.12.043

37. Chenlo Barros R, Gómez Rodríguez BJ, García Soidan FJ, Clavería Fontán A. Prevalencia de enfermedades dermatológicas en atención primaria. Med Fam SEMERGEN. 2021;47(7):434-40. DOI: 10.1016/j.semerg.2021.01.012

38. Verhoeven EWM, Kraaimaat FW, van Weel C, van de Kerkhof PCM, Duller P, van der Valk PGM, et al. Skin diseases in family medicine: prevalence and health care use. Ann Fam Med. 2008;6(4):349-54. DOI: 10.1370/afm.861

39. Coombes-Perez C, Arribas-Garcia P, Llerena-Delgado A, Del Solar-Chacaltana MA, Carrillo-Larco RM. Skin Lesions and Personal Protective Equipment in Health Care Workers From Lima, Peru, During the COVID-19 Pandemic: A Cross-sectional Study. Open Forum Infect Dis. 2025;12(9):ofaf509. DOI: 10.1093/ofid/ofaf509

40. Seth D, Cheldize K, Brown D, Freeman EE. Global Burden of Skin Disease: Inequities and Innovations. Curr Dermatol Rep. 2017;6(3):204-10. DOI: 10.1007/s13671-017-0192-7

41. Ko CJ, Braverman I, Sidlow R, Lowenstein EJ. Visual perception, cognition, and error in dermatologic diagnosis: Key cognitive principles. J Am Acad Dermatol. 2019;81(6):1227-34. DOI: 10.1016/j.jaad.2018.10.082

42. Dugonik B, Golob M, Marhl M, Dugonik A. Optimizing Digital Image Quality for Improved Skin Cancer Detection. J Imaging. 2025;11(4):107. DOI: 10.3390/jimaging11040107

43. Parekh P, Oyeleke R, Vishwanath T. The Depth Estimation and Visualization of Dermatological Lesions: Development and Usability Study. JMIR Dermatol. 2024;7:e59839. DOI: 10.2196/59839

44. Surya A, K D, Sekar M, Preetham C, Sudha S, A S. Clinical Photography in Dermatology: A Cross-Sectional Analysis From the Patient's Perspective in a Tertiary Care Center in Tamil Nadu, India. Cureus. 2025;17(7):e88112. DOI: 10.7759/cureus.88112

45. Norman GR, Rosenthal D, Brooks LR, Allen SW, Muzzin LJ. The development of expertise in dermatology. Arch Dermatol. 1989;125(8):1063-8.

46. Doyle C, Tobin A-M, Foley T. 'I'm definitely more confident now': a qualitative study exploring how GP trainees learn about dermatology. Skin Health Dis. 2025;5(4):273-80. DOI: 10.1093/skinhd/vzaf043

47. Loveday T, Wiggins M, Festa M, Schell D, Twigg D. Pattern Recognition as an Indicator of Diagnostic Expertise. En: Latorre Carmona P, Sánchez JS, Fred ALN, editores. Pattern Recognition - Applications and Methods. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. (Advances in Intelligent Systems and Computing; vol. 204). DOI: 10.1007/978-3-642-36530-0_1

48. van de Wiel MWJ, Van den Bossche P, Janssen S, Jossberger H. Exploring deliberate practice in medicine: how do physicians learn in the workplace? Adv Health Sci Educ Theory Pract. 2011;16(1):81-95. DOI: 10.1007/s10459-010-9246-3

49. Anders Ericsson K. Deliberate Practice and Acquisition of Expert Performance: A General Overview. Acad Emerg Med. 2008;15(11):988-94. DOI: 10.1111/j.1553-2712.2008.00227.x

50. McGaghie WC, Kristopaitis T. Deliberate practice and mastery learning: origins of expert medical performance. En: Cleland J, Durning SJ, editores. Researching Medical Education. 1.ªed. Wiley; 2022. DOI: 10.1002/9781119839446.ch28

51. Sung T-C, Shih H-I, Kawaguchi T, Chi C-H, Hsu H-C. Tracking Clinical Competency Growth: A Longitudinal Study of Medical Students in a Multidisciplinary Emergency Department Internship Program. J Multidiscip Healthc. 2025;18:3877-90. DOI: 10.2147/JMDH.S530887

52. Pragya A N, Juhi S, Jinal T, Dharmesh P. Diagnostic criteria in dermatology. Indian J Dermatol Venereol Leprol. 2023;89(5):771-9. DOI: 10.25259/IJDVL_24_2022

53. Williams HC, Burden-Teh E. On the definition of dermatological disease. Part 2: approaches for defining dermatological diseases. Clin Exp Dermatol. 2022;47(10):1812-9. DOI: 10.1111/ced.15278

54. Gulzar Y, Agarwal S, Soomro S, Kandpal M, Turaev S, Onn CW, et al. Next-generation approach to skin disorder prediction employing hybrid deep transfer learning. Front Big Data. 2025;8:1503883. DOI: 10.3389/fdata.2025.1503883

55. Benelli C, Roscetti E, Dal Pozzo V. Reproducibility of the clinical criteria (ABCDE rule) and dermatoscopic features (7FFM) for the diagnosis of malignant melanoma. Eur J Dermatol EJD. 2001;11(3):234-9.

56. Kashetsky N, Mar K, Liu C, Rivers JK, Mukovozov I. Photography in dermatology - a scoping review: Practices, skin of color, patient preferences, and medical-legal considerations. JDDG J Dtsch Dermatol Ges. 2023;21(10):1102-7. DOI: 10.1111/ddg.15129

57. Abbasi NR, Shaw HM, Rigel DS, Friedman RJ, McCarthy WH, Osman I, et al. Early Diagnosis of Cutaneous Melanoma: Revisiting the ABCD Criteria. JAMA. 2004;292(22):2771. DOI: 10.1001/jama.292.22.2771

58. Savelyev DS, Gorodkov SYu, Goremykin IV. Standardization of color measurement in the medical photography in clinical practice. Russ J Pediatr Surg. 2024;28(5):460-71. DOI: 10.17816/ps803

59. Adya KA. Clinical and Demographic Characteristics of Dermatoses with Bimodal Distribution Pattern: A Concise Review. Indian Dermatol Online J. 2025;16(5):717-24. DOI: 10.4103/idoj.idoj_1288_24

60. Holmes Z, Courtney A, Lincoln M, Weller R. Rash Morphology as a Predictor of COVID-19 Severity: A Systematic Review of the Cutaneous Manifestations of COVID-19. Skin Health Dis. 2022;2(3):ski2.120. DOI: 10.1002/ski2.120

61. AlFada M, Alotaibi H, Alsharif S, Alani AH, Andrade-Miranda A, Montesinos Guevara C, et al. Systematic review, methodological appraisal, and recommendation mapping of clinical practice guidelines for managing patients with Stevens-Johnson syndrome and toxic epidermal necrolysis. J Dermatol Treat. 2025;36(1):2467751. DOI: 10.1080/09546634.2025.2467751

62. Tamburus ALP, Marcello VZ, Baleotti LR, Oliveira PHCD, Amorim SAD, Braz JCM, et al. Desafíos de la concordancia diagnóstica en dermatología: un análisis

interinstitucional. Cogitare Enferm. 2025;30:e98057pt. DOI: 10.1590/ce.v30i0.98057es

63. Mevorach L, Farcomeni A, Pellacani G, Cantisani C. A Comparison of Skin Lesions' Diagnoses Between AI-Based Image Classification, an Expert Dermatologist, and a Non-Expert. Diagnostics. 2025;15(9):1115. DOI: 10.3390/diagnostics15091115

64. Noor MN, Haneef F, Ashraf I, Masud M. Enhanced Skin Disease Classification via Dataset Refinement and Attention-Based Vision Approach. Bioengineering. 2025;12(3):275. DOI: 10.3390/bioengineering12030275

65. Uwishema O, Ghezzawi M, Charbel N, Alawieh S, Roy S, Wojtara M, et al. Diagnostic performance of artificial intelligence for dermatological conditions: a systematic review focused on low- and middle-income countries to address resource constraints and improve access to specialist care. Int J Emerg Med. 2025;18(1):172. DOI: 10.1186/s12245-025-00975-4

66. Tweed M, Wilkinson T. Diagnostic testing and educational assessment. Clin Teach. 2012;9(5):299-303. DOI: 10.1111/j.1743-498X.2012.00567.x

67. Fink MC, Reitmeier V, Stadler M, Siebeck M, Fischer F, Fischer MR. Assessment of Diagnostic Competences With Standardized Patients Versus Virtual Patients: Experimental Study in the Context of History Taking. J Med Internet Res. 2021;23(3):e21196. DOI: 10.2196/21196

68. Smith SK, Benbenek MM, Bakker CJ, Bockwoldt D. Scoping review: Diagnostic reasoning as a component of clinical reasoning in the U.S. primary care nurse practitioner education. J Adv Nurs. 2022;78(12):3869-96. DOI: 10.1111/jan.15414

69. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol. 2022;75(1):25-36. DOI: 10.4097/kja.21209

70. Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. Npj Digit Med. 2021;4(1):65. DOI: 10.1038/s41746-021-00438-z

71. Agraz M, Mantzoros C, Karniadakis GE. Shimosawa T, editor. ChatGPT-Enhanced ROC Analysis (CERA): A shiny web tool for finding optimal cutoff points in biomarker analysis. PLOS ONE. 2024;19(4):e0289141. DOI: 10.1371/journal.pone.0289141

72. Aguilar-Ruiz JS, Michalak M. Classification performance assessment for imbalanced multiclass data. Sci Rep. 2024;14(1):10759. DOI: 10.1038/s41598-024-61365-z

73. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 2020;21(1):6. DOI: 10.1186/s12864-019-6413-7

74. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006;27(8):861-74. DOI: 10.1016/j.patrec.2005.10.010

75. Albitres-Flores L, Pisfil-Farroñay YA, Guillen-Macedo K, Niño-Garcia R, Alarcon-Ruiz CA. Percepción de los internos sobre la suspensión del internado

médico durante la cuarentena por la COVID-19. Rev Peru Med Exp Salud Pública. 2020;37(3):504-9. DOI: 10.17843/rpmesp.2020.373.5729

76. Lewis TP, Roder-DeWan S, Malata A, Ndiaye Y, Kruk ME. Clinical performance among recent graduates in nine low- and middle-income countries. Trop Med Int Health. 2019;24(5):620-35. DOI: 10.1111/tmi.13224

77. Lovén M, Huilaja L, Paananen M, Torkki P. The integration of dermatology experts into primary care to assess and treat patients with skin lesions is cost-effective: A quasi-experimental study. J Eur Acad Dermatol Venereol. 2025;39(9):1666-74. DOI: 10.1111/jdv.20451

78. Lowell BA, Froelich CW, Federman DG, Kirsner RS. Dermatology in primary care: Prevalence and patient disposition. J Am Acad Dermatol. 2001;45(2):250-5. DOI: 10.1067/mjd.2001.114598

79. Le Roux E, Edwards PJ, Sanderson E, Barnes RK, Ridd MJ. The content and conduct of GP consultations for dermatology problems: a cross-sectional study. Br J Gen Pract. 2020;70(699):e723-30. DOI: 10.3399/bjgp20X712577

80. Rimoin L, Altieri L, Craft N, Krasne S, Kellman PJ. Training pattern recognition of skin lesion morphology, configuration, and distribution. J Am Acad Dermatol. 2015;72(3):489-95. DOI: 10.1016/j.jaad.2014.11.016

81. Scott BL, Barker B, Abraham R, Wickless HW. Integration of Dermatology-Focused Physical Diagnosis Rounds and Case-Based Learning within the Internal Medicine Medical Student Clerkship. J Med Educ Curric Dev. 2016;3:JMECD.S40417. DOI: 10.4137/JMECD.S40417

82. Tong LZ, Desai RM, Olsen R, Davis M. The Pathophysiology, Diagnosis and Management of Chronic Inflammatory Skin Diseases. Discov Med. 2024;36(189):1933. DOI: 10.24976/Discov.Med.202436189.180

83. Clarke P. Benign pigmented skin lesions. Aust J Gen Pract. 2019;48(6):364-7. DOI: 10.31128/AJGP-12-18-4802

84. Greco MJ, Bhutta BS. Seborrheic Keratosis. En: StatPearls. [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 [citado 22 de noviembre de 2025]. Recuperado: http://www.ncbi.nlm.nih.gov/books/NBK545285/

85. McMahon DE, Schuetz AN, Kovarik CL. Emerging infectious diseases of the skin: a review of clinical and histologic findings. Hum Pathol. 2023;140:196-213. DOI: 10.1016/j.humpath.2023.07.003

86. World Health Organization. International Classification of Diseases (ICD) [Internet]. 2021. International Classification of Diseases (ICD) [citado 13 de noviembre de 2025]. Recuperado: https://www.who.int/standards/classifications/classification-of-diseases

87. Salava A, Salmela V. Diagnostic errors during perceptual learning in dermatology: a prospective cohort study of Finnish undergraduate students. Clin Exp Dermatol. 2024;49(8):866-74. DOI: 10.1093/ced/llae063

88. Reddy S, Shaheed A, Patel R. Artificial Intelligence in Dermoscopy: Enhancing Diagnosis to Distinguish Benign and Malignant Skin Lesions. Cureus. 2024; DOI: 10.7759/cureus.54656

89. Das A, Agarwal V, Shetty NP. Comparative analysis of multimodal architectures for effective skin lesion detection using clinical and image data. Front Artif Intell. 2025;8:1608837. DOI: 10.3389/frai.2025.1608837

90. Hernández Nieto H. Contributions to statistical analysis: sensibility (estability and consistency) of different coefficients of relative variance and the coefficient of proportional variance (Cvp) : the coefficient of content validity (Ccv) and the Kappa coefficient (K), to determine content validity according to the technique of experts panel. Mérida, Venezuela: Universidad de Los Andes; 2002.

7. ANEXOS

7.1. Anexo 1: Matriz de consistencia

PREGUNTA DE INVESTIGACIÓN	OBJETIVOS	HIPÓTESIS	VARIABLES	METODOLOGÍA	ANÁLISIS
Pregunta general: ¿Cuál es el nivel de precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM en 2025? **Preguntas específicas:** ¿Cuál es la precisión diagnóstica global y específica por cada una de las 15 patologías dermatológicas evaluadas mediante imágenes clínicas?	**Objetivo general:** Evaluar la precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM en 2025. **Objetivos específicos:** Determinar la precisión diagnóstica global y por cada una de las 15 patologías dermatológicas evaluadas. Identificar los patrones de confusión diagnóstica más frecuentes entre las patologías	**Hipótesis general** El nivel de precisión diagnóstica en dermatología mediante imágenes clínicas en los egresados de medicina de la UNMSM del 2025 es de nivel intermedio o regular, reflejando las brechas en el entrenamiento diagnóstico visual reportadas en la literatura. **Hipótesis específicas** **Hipótesis específica 1:** Existen diferencias estadísticamente significativas en la precisión diagnóstica entre las distintas categorías clínicas de patologías dermatológicas evaluadas (inflamatorias, infecciosas, pigmentadas benignas y premalignas/malignas) en los	**Variable de Estudio:** Precisión Diagnóstica Dermatológica. **Variable de Análisis de Error:** Patrón de Confusión Diagnóstica. **Variables Sociodemográficas y Académicas:** - Edad - Sexo biológico - Sede de internado	**Tipo de investigación:** Estudio observacional transversal descriptivo-comparativo. **Población**: Todos los egresados de medicina de la UNMSM que culminaron el internado médico obligatorio y egresaron en el año académico 2025 (N=185). **Muestra**: Población censal accesible. Se invitó a participar a la totalidad de los 185 egresados; se excluyó a 30 que habían participado en el estudio piloto del	Procesamiento y análisis de datos en SPSS v26 y Python 3.x (pandas, numpy, seaborn/matplotlib) para garantizar análisis reproducibles. Estadística descriptiva de las características de los egresados (frecuencias y porcentajes en variables categóricas, medidas de tendencia central y dispersión en variables numéricas). Cálculo de precisión diagnóstica global, por cada una de las 15 patologías y por categorías clínicas principales, expresada en porcentajes. Construcción y análisis de

¿Existen diferencias en la precisión diagnóstica entre las categorías de enfermedades dermatológicas evaluadas (inflamatorias, infecciosas, pigmentadas benignas y premalignas/malignas)? ¿Cuáles son los pares de dermatosis más frecuentemente confundidos y qué patrones de error diagnóstico se identifican mediante el análisis de la matriz de confusión?	dermatológicas evaluadas. Comparar la precisión diagnóstica entre las categorías clínicas de patologías dermatológicas evaluadas. Comparar la precisión diagnóstica según las características sociodemográficas y académicas (sexo, edad, sede de internado) de los participantes.	egresados de medicina de la UNMSM del 2025. **Hipótesis específica 2:** Existen diferencias estadísticamente significativas en la precisión diagnóstica según las características sociodemográficas y académicas (sexo, edad y sede de internado) de los egresados de medicina de la UNMSM del 2025.		instrumento, y la muestra final estuvo constituida por 113 egresados que completaron el cuestionario. **Instrumento**: Cuestionario de Precisión Diagnóstica Dermatológica mediante Imágenes Clínicas (CPDD-IC), validado mediante juicio de expertos para contenido y sometido a valoración psicométrica en un estudio piloto previo.	una matriz de confusión 15×15 para identificar pares de patologías frecuentemente confundidos y describir patrones sistemáticos de error. Aplicación de Q de Cochran, prueba de Friedman (con post-hoc de Wilcoxon) y pruebas t/ANOVA o Mann–Whitney/Kruskal–Wallis, así como correlación de Pearson/Spearman, para evaluar diferencias en precisión diagnóstica entre patologías, categorías clínicas y subgrupos (sexo, sede, edad), con nivel de significancia de 0,05.

7.2. Anexo 2: Resolución decanal de la aprobación del proyecto de tesis



UNMSM

Firmado digitalmente por DEL CARMEN SARA Jose Carlos FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 13.02.2026 10:17:25 -05:00

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú. Decana de América
FACULTAD DE MEDICINA
DECANATO

Lima, 13 de Febrero del 2026

RESOLUCIÓN DECANAL N° 000744-2026-D-FM/UNMSM

Visto el expediente digital UNMSM-20250147637, de fecha 17 de diciembre de 2025 de la Facultad de Medicina, sobre aprobación de Proyecto de tesis

CONSIDERANDO:

Que, mediante Resolución Decanal N° 002373-2023-D-FM/UNMSM de fecha 19 de junio de 2023 ratificada con Resolución Rectoral N° 003396-2024 de fecha 04 de marzo del 2024 se aprueba el Reglamento para la Elaboración de Tesis para optar el Título Profesional en las Escuelas Profesionales de la Facultad de Medicina, que en su Capítulo V. DE LA PLANIFICACION, Art. 8: establece que: "La tesis (T) podrá ser individual o grupal (máximo tres), compartiéndose las responsabilidades de la elaboración." así mismo, en su Capítulo VI: Del Asesoramiento aprobación e inscripción del proyecto de tesis: Art. 23 establece que: "Con la aprobación del proyecto por parte del Comité de Investigación correspondiente y, de un Comité de Ética en Investigación, la Dirección de la EP solicitara al Vicedecanato Académico la Resolución Decanal de la aprobación del proyecto, y el tesista podrá iniciar la ejecución de su proyecto de tesis.";

Que, mediante Oficio N°000210-2026-EPMH-FM/UNMSM, el Director (e) de la Escuela Profesional de Medicina Humana; eleva el Informe de don José Percy Amado Tineo, docente asociado del Departamento Académico de Medicina Humana y Jurado informante del Proyecto de Tesis titulado: "Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025", presentado por el egresado Sr. Walter Allccarima Ramos, con código de matrícula 18010285, para optar el Título Profesional de Médico Cirujano; informa que el Proyecto de Tesis mencionado se encuentra APTO para ser ejecutado; por lo que, solicita autorizar la emisión de la Resolución de Decanato respectiva incluyendo el nombre del asesor de la tesis Mg. Gadwyn Sanchez Felix, con código 081884, docente asociado del Departamento Académico de Medicina Humana; y,

Estando a lo establecido por el Estatuto de la Universidad Nacional Mayor de San Marcos y las atribuciones conferidas por la Ley Universitaria N°30220;

SE RESUELVE:

1° Aprobar el Proyecto de Tesis, según detalle:

Egresado: **Walter Allccarima Ramos** **Código de matrícula N° 18010285** **E.P. de Medicina Humana**	**Título del Proyecto de Tesis:** **Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025**
Asesor: **Mg. Gadwyn Sanchez Felix** **Código: 081884**	

2° Encargar a la Escuela Profesional de Medicina Humana el cumplimiento de la presente resolución.

Regístrese, comuníquese, archívese.

UNMSM

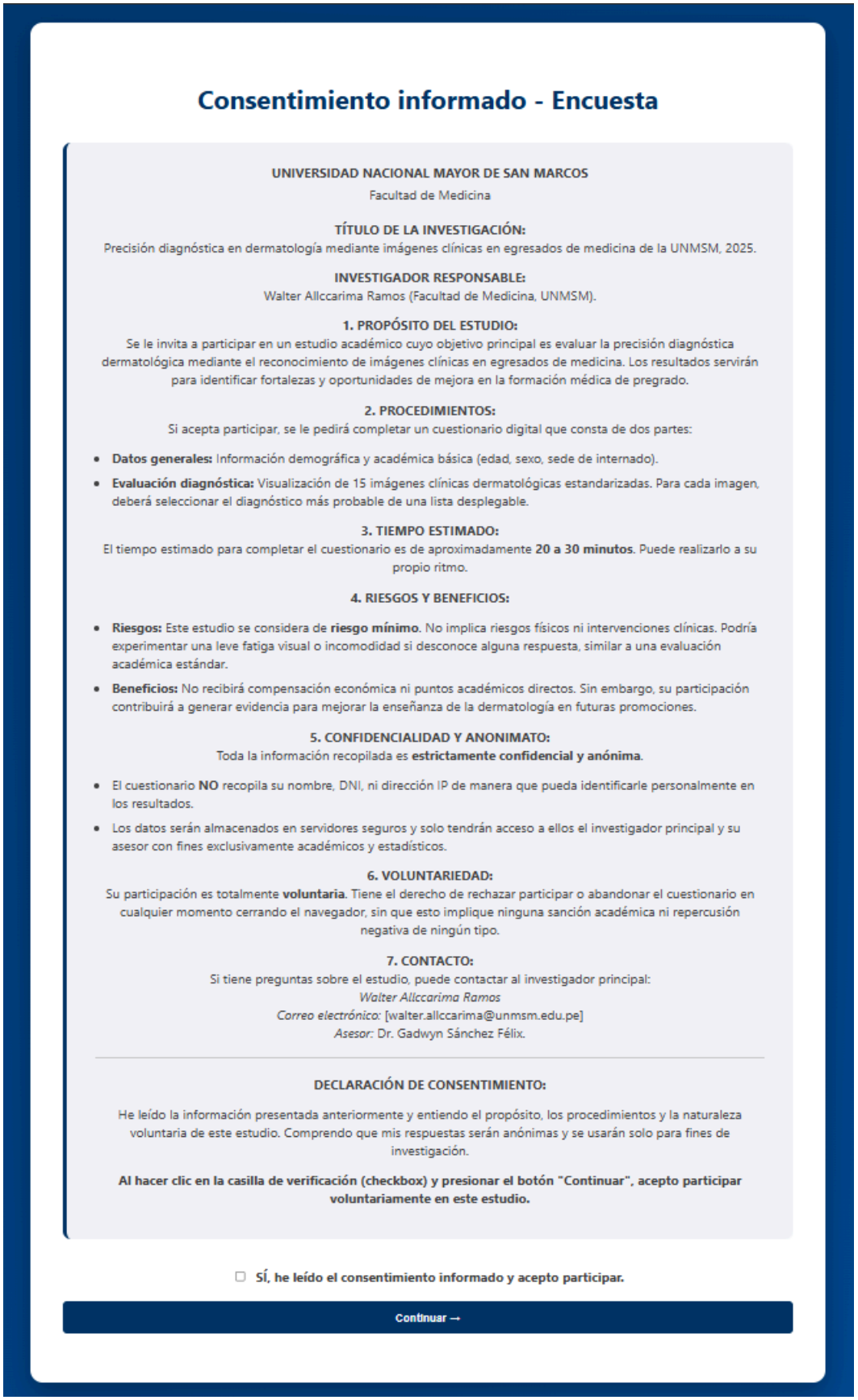
Dra. EMMA FELICIA SALAZAR SALVATIERRA
Vicedecana Académica (e)

Dr. JOSÉ CARLOS DEL CARMEN SARA
Decano

Firmado digitalmente por SALAZARJDS/mmb
SALVATIERRA Emma Felicia FAU
20148092282 soft
Motivo: Doy V°B°
Fecha: 12.02.2026 20:34:16 -05:00

Esta es una copia auténtica imprimible de un documento electrónico archivado por la Universidad Nacional Mayor de San Marcos, aplicando lo dispuesto por el Art. 25 de D.S. 070-2013-PCM y la Tercera Disposición Complementaria Final del D.S. 026- 2016-PCM. Su autenticidad e integridad pueden ser contrastadas a través de la siguiente dirección web: **https://spsgd.unmsm.edu.pe/verifica/inicio.do** e ingresando el siguiente código de verificación: **RXQMMSQ**

7.3. Anexo 3: Consentimiento informado



Consentimiento informado - Encuesta

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Facultad de Medicina

TÍTULO DE LA INVESTIGACIÓN:

Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025.

INVESTIGADOR RESPONSABLE:

Walter Allccarima Ramos (Facultad de Medicina, UNMSM).

1. PROPÓSITO DEL ESTUDIO:

Se le invita a participar en un estudio académico cuyo objetivo principal es evaluar la precisión diagnóstica dermatológica mediante el reconocimiento de imágenes clínicas en egresados de medicina. Los resultados servirán para identificar fortalezas y oportunidades de mejora en la formación médica de pregrado.

2. PROCEDIMIENTOS:

Si acepta participar, se le pedirá completar un cuestionario digital que consta de dos partes:

- **Datos generales:** Información demográfica y académica básica (edad, sexo, sede de internado).
- **Evaluación diagnóstica:** Visualización de 15 imágenes clínicas dermatológicas estandarizadas. Para cada imagen, deberá seleccionar el diagnóstico más probable de una lista desplegable.

3. TIEMPO ESTIMADO:

El tiempo estimado para completar el cuestionario es de aproximadamente **20 a 30 minutos**. Puede realizarlo a su propio ritmo.

4. RIESGOS Y BENEFICIOS:

- **Riesgos:** Este estudio se considera de **riesgo mínimo**. No implica riesgos físicos ni intervenciones clínicas. Podría experimentar una leve fatiga visual o incomodidad si desconoce alguna respuesta, similar a una evaluación académica estándar.
- **Beneficios:** No recibirá compensación económica ni puntos académicos directos. Sin embargo, su participación contribuirá a generar evidencia para mejorar la enseñanza de la dermatología en futuras promociones.

5. CONFIDENCIALIDAD Y ANONIMATO:

Toda la información recopilada es **estrictamente confidencial y anónima**.

- El cuestionario **NO** recopila su nombre, DNI, ni dirección IP de manera que pueda identificarle personalmente en los resultados.
- Los datos serán almacenados en servidores seguros y solo tendrán acceso a ellos el investigador principal y su asesor con fines exclusivamente académicos y estadísticos.

6. VOLUNTARIEDAD:

Su participación es totalmente **voluntaria**. Tiene el derecho de rechazar participar o abandonar el cuestionario en cualquier momento cerrando el navegador, sin que esto implique ninguna sanción académica ni repercusión negativa de ningún tipo.

7. CONTACTO:

Si tiene preguntas sobre el estudio, puede contactar al investigador principal:

Walter Allccarima Ramos

Correo electrónico: [walter.allccarima@unmsm.edu.pe]

Asesor: Dr. Gadwyn Sánchez Félix.

DECLARACIÓN DE CONSENTIMIENTO:

He leído la información presentada anteriormente y entiendo el propósito, los procedimientos y la naturaleza voluntaria de este estudio. Comprendo que mis respuestas serán anónimas y se usarán solo para fines de investigación.

Al hacer clic en la casilla de verificación (checkbox) y presionar el botón "Continuar", acepto participar voluntariamente en este estudio.

☐ **SÍ, he leído el consentimiento informado y acepto participar.**

Continuar →

7.4. Anexo 4: Cuestionario de Precisión Diagnóstica Dermatológica mediante Imágenes Clínicas (CPDD-IC)

El enlace para la visualización del cuestionario es el siguiente: https://cuestionario-dermatologia.vercel.app





Pregunta 1 de 15



Selecciona el diagnóstico:

-- Selecciona una opción --

Siguiente →

Pregunta 4 de 15



Selecciona el diagnóstico:

-- Selecciona una opción --

← Anterior

Siguiente →

7.5. Anexo 5: Evaluación Likert de 5 puntos

VALIDACIÓN DE EXPERTOS

Cuestionario de Precisión Diagnóstica Dermatológica - CPDD-IC

Página 3 de 19

3. INSTRUCCIONES PARA LA EVALUACIÓN

Escala de Calificación (1-5):

Representatividad Diagnóstica:

- **1-2:** No representativa
- **3:** Moderadamente representativa
- **4-5:** Representativa o altamente representativa

Calidad Técnica Fotográfica:

- **1-2:** Pobre/regular
- **3:** Aceptable
- **4-5:** Buena/excelente

Ausencia de Sesgo Visual:

- **1-2:** Presencia evidente de pistas
- **3:** Pocas pistas sesgadas
- **4-5:** Sin pistas no intencionadas

← Anterior | Siguiente →

Califique cada criterio de 1 (inaceptable) a 5 (excelente)

Imagen 4: Impétigo no bulloso



	1	2	3	4	5
Representatividad diagnóstica					
Calidad técnica fotográfica					
Ausencia de sesgo visual					

7.6. Anexo 6: Validación por juicio de expertos



Validación de Expertos - Instrumento CPDD-IC

Panel de 5 Dermatólogos Certificados - Fase de Validación de Contenido

Noviembre 2025

Experto 1: Paredes L. | Experto 2: Espinoza E. | Experto 3: Paredes A. | Experto 4: Kossuth C. | Experto 5: Palacios C.

Guido Oswaldo Aurelio Paredes Llerena

CMP: 16283

RNE: 8063

Institución: Consultorio Privado, **Experto Validador N°1**

Años de experiencia: 42

Fecha de evaluación: 09/11/2025

Representatividad Diagnóstica	Calidad Técnica Fotográfica	Ausencia de Sesgo Visual	PROMEDIO GENERAL
5.00/5.0	5.00/5.0	5.00/5.0	5.00/5.0

Evaluación Detallada por Patología (15 Imágenes)

#	Patología Dermatológica	Representatividad	Calidad Técnica	Ausencia Sesgo	Promedio
1	Acné vulgar	5	5	5	5.00
2	Carcinoma basocelular	5	5	5	5.00
3	Dermatitis atópica	5	5	5	5.00
4	Dermatitis de contacto alérgica	5	5	5	5.00
5	Impétigo no bulloso	5	5	5	5.00
6	Lentigo solar	5	5	5	5.00
7	Melanoma	5	5	5	5.00
8	Molusco contagioso	5	5	5	5.00
9	Nevo melanocítico común	5	5	5	5.00
10	Psoriasis	5	5	5	5.00
11	Queratosis actínica	5	5	5	5.00
12	Queratosis seborreica	5	5	5	5.00
13	Rosácea	5	5	5	5.00
14	Tiña corporis	5	5	5	5.00
15	Verruga vulgar	5	5	5	5.00

Validación de Expertos - Instrumento CPDD-IC

Panel de 5 Dermatólogos Certificados - Fase de Validación de Contenido

Noviembre 2025

Experto 1: Paredes L. | **Experto 2: Espinoza E.** | Experto 3: Paredes A. | Experto 4: Kossuth C. | Experto 5: Palacios C.

Renzo Bruno Espinoza Escudero

CMP: 85527

RNE: 49388

Institución: Clínica Privada, Cons

Experto Validador N°2

Años de experiencia: 5

Fecha de evaluación: 09/11/2025

Representatividad Diagnóstica	Calidad Técnica Fotográfica	Ausencia de Sesgo Visual	PROMEDIO GENERAL
4.73/5.0	5.00/5.0	5.00/5.0	4.91/5.0

Evaluación Detallada por Patología (15 Imágenes)

#	Patología Dermatológica	Representatividad	Calidad Técnica	Ausencia Sesgo	Promedio
1	Acné vulgar	3	5	5	4.33
2	Carcinoma basocelular	5	5	5	5.00
3	Dermatitis atópica	5	5	5	5.00
4	Dermatitis de contacto alérgica	5	5	5	5.00
5	Impétigo no bulloso	5	5	5	5.00
6	Lentigo solar	5	5	5	5.00
7	Melanoma	5	5	5	5.00
8	Molusco contagioso	5	5	5	5.00
9	Nevo melanocítico común	5	5	5	5.00
10	Psoriasis	5	5	5	5.00
11	Queratosis actínica	4	5	5	4.67
12	Queratosis seborreica	5	5	5	5.00
13	Rosácea	5	5	5	5.00
14	Tiña corporis	5	5	5	5.00
15	Verruga vulgar	5	5	5	5.00

Validación de Expertos - Instrumento CPDD-IC

Panel de 5 Dermatólogos Certificados - Fase de Validación de Contenido

Noviembre 2025

Experto 1: Paredes L. | Experto 2: Espinoza E. | **Experto 3: Paredes A.** | Experto 4: Kossuth C. | Experto 5: Palacios C.

Antonio Arnaldo Paredes Arcos

CMP: 13820

RNE: 5800

Institución: Clínica Privada

Años de experiencia: 39

Fecha de evaluación: 11/11/2025

Experto Validador N°3

Representatividad Diagnóstica	Calidad Técnica Fotográfica	Ausencia de Sesgo Visual	PROMEDIO GENERAL
4.60/5.0	4.67/5.0	4.67/5.0	4.64/5.0

Evaluación Detallada por Patología (15 Imágenes)

#	Patología Dermatológica	Representatividad	Calidad Técnica	Ausencia Sesgo	Promedio
1	Acné vulgar	4	4	5	4.33
2	Carcinoma basocelular	5	5	5	5.00
3	Dermatitis atópica	5	5	5	5.00
4	Dermatitis de contacto alérgica	5	5	5	5.00
5	Impétigo no bulloso	5	5	5	5.00
6	Lentigo solar	5	5	5	5.00
7	Melanoma	5	5	5	5.00
8	Molusco contagioso	5	5	5	5.00
9	Nevo melanocítico común	5	5	5	5.00
10	Psoriasis	5	5	5	5.00
11	Queratosis actínica	4	4	4	4.00
12	Queratosis seborreica	4	4	5	4.33
13	Rosácea	5	5	5	5.00
14	Tiña corporis	4	4	4	4.00
15	Verruga vulgar	5	5	5	5.00

Validación de Expertos - Instrumento CPDD-IC

Panel de 5 Dermatólogos Certificados - Fase de Validación de Contenido

Noviembre 2025

Experto 1: Paredes L. | Experto 2: Espinoza E. | Experto 3: Paredes A. | **Experto 4: Kossuth C.** | Experto 5: Palacios C.

Stefano Kossuth Cabrejos

CMP: 67137

RNE: 36931

Institución: Clínica Privada

Años de experiencia: 7

Fecha de evaluación: 14/11/2025

Experto Validador N°4

Representatividad Diagnóstica	Calidad Técnica Fotográfica	Ausencia de Sesgo Visual	PROMEDIO GENERAL
3.67/5.0	3.67/5.0	3.47/5.0	3.60/5.0

Evaluación Detallada por Patología (15 Imágenes)

#	Patología Dermatológica	Representatividad	Calidad Técnica	Ausencia Sesgo	Promedio
1	Acné vulgar	4	4	4	4.00
2	Carcinoma basocelular	4	4	4	4.00
3	Dermatitis atópica	4	4	4	4.00
4	Dermatitis de contacto alérgica	4	4	4	4.00
5	Impétigo no bulloso	4	4	4	4.00
6	Lentigo solar	4	4	4	4.00
7	Melanoma	3	3	2	2.67
8	Molusco contagioso	3	3	4	3.33
9	Nevo melanocítico común	4	4	3	3.67
10	Psoriasis	3	3	2	2.67
11	Queratosis actínica	4	4	4	4.00
12	Queratosis seborreica	2	4	2	2.67
13	Rosácea	2	4	3	3.00
14	Tiña corporis	4	4	4	4.00
15	Verruga vulgar	4	4	4	4.00

Validación de Expertos - Instrumento CPDD-IC

Panel de 5 Dermatólogos Certificados - Fase de Validación de Contenido

Noviembre 2025

Experto 1: Paredes L. | Experto 2: Espinoza E. | Experto 3: Paredes A. | Experto 4: Kossuth C. | **Experto 5: Palacios C.**

Christian Andre Palacios Curay

Experto Validador N°5

- CMP: 74389
- RNE: 40483
- Institución: Clínica Privada, Con
- Años de experiencia: 5
- Fecha de evaluación: 14/11/2025

Representatividad Diagnóstica	Calidad Técnica Fotográfica	Ausencia de Sesgo Visual	PROMEDIO GENERAL
5.00/5.0	5.00/5.0	5.00/5.0	5.00/5.0

Evaluación Detallada por Patología (15 Imágenes)

#	Patología Dermatológica	Representatividad	Calidad Técnica	Ausencia Sesgo	Promedio
1	Acné vulgar	5	5	5	5.00
2	Carcinoma basocelular	5	5	5	5.00
3	Dermatitis atópica	5	5	5	5.00
4	Dermatitis de contacto alérgica	5	5	5	5.00
5	Impétigo no bulloso	5	5	5	5.00
6	Lentigo solar	5	5	5	5.00
7	Melanoma	5	5	5	5.00
8	Molusco contagioso	5	5	5	5.00
9	Nevo melanocítico común	5	5	5	5.00
10	Psoriasis	5	5	5	5.00
11	Queratosis actínica	5	5	5	5.00
12	Queratosis seborreica	5	5	5	5.00
13	Rosácea	5	5	5	5.00
14	Tiña corporis	5	5	5	5.00
15	Verruga vulgar	5	5	5	5.00

Cada experto evaluó de forma independiente las 15 imágenes clínicas fotográficas incluidas en el instrumento, calificando tres criterios mediante escala tipo Likert de 1 (inadecuado) a 5 (muy adecuado):

- **Representatividad diagnóstica (1-5):** Grado en que la imagen ejemplifica las características típicas y diagnósticas de la patología, observable por el médico general.
    - 1 = No es representativa de la patología; carece de características diagnósticas clave
    - 2 = Débilmente representativa; presenta algunos rasgos pero de forma atípica o ambigua
    - 3 = Moderadamente representativa; presenta características reconocibles aunque no típicas
    - 4 = Altamente representativa; muestra características diagnósticas claras y típicas
    - 5 = Altamente representativa de la patología; excelentes características diagnósticas definitorias

Justificación metodológica: La calidad técnica es determinante en el diagnóstico visual. Imágenes de baja calidad pueden inducir a error independientemente de su contenido.

- **Calidad técnica fotográfica (1-5):** Adecuación técnica de la imagen para propósitos diagnósticos-educativos.
    - 1 = Pobre; baja resolución, iluminación inadecuada, enfoque deficiente, colores distorsionados
    - 2 = Aceptable pero deficiente; algunos problemas técnicos que limitan la visualización
    - 3 = Adecuada; calidad técnica media, aceptable para educación
    - 4 = Buena; calidad técnica satisfactoria con mínimos defectos
    - 5 = Excelente; alta resolución, iluminación óptima, enfoque nítido, colores precisos, sin artefactos

Justificación metodológica: La calidad técnica es determinante en el diagnóstico visual. Imágenes de baja calidad pueden inducir a error independientemente de su contenido.

- **Ausencia de sesgo visual (1-5):** Grado en que la imagen evita proporcionar pistas diagnósticas no pertinentes o que faciliten la adivinación.
    - 1 = Presenta pistas diagnósticas evidentes; prácticamente obliga la respuesta correcta sin análisis
    - 2 = Presenta varias pistas sesgadas; facilita significativamente la respuesta
    - 3 = Presenta algunas pistas sesgadas; moderadamente facilitador
    - 4 = Presenta pistas mínimas; requiere análisis clínico substantivo
    - 5 = Sin pistas sesgadas; requiere análisis clínico riguroso sin facilitadores

Justificación educativa: Para que la evaluación sea válida, debe medir conocimiento real, no la capacidad de percibir señales incidentales.

Índice de Validez de Contenido (IVC)

Para cuantificar el acuerdo de los expertos, se calculó el Índice de Validez de Contenido (IVC) propuesto por Lawshe (1975) y modificado por Lynn (1986). La fórmula del IVC es:

$$IVC = \frac{N_{4-5}}{N_{Total}}$$

Donde:

- N_{4-5} = Número de expertos que otorgan puntaje igual o superior a 4 (puntos 4 ó 5)
- N Total = Número total de expertos evaluadores (n=5)

El estándar internacional de validez de contenido establece que se debe aceptar una imagen cuando el IVC ≥ 0.80 en cada uno de los tres criterios. Esto

equivale al consenso de al menos 4 de 5 expertos considerando la respuesta como "buena o excelente" (puntuación ≥ 4).

Se consideró aceptable una imagen cuando simultáneamente:

- IVC (Representatividad) ≥ 0.80
- IVC (Calidad técnica) ≥ 0.80
- IVC (Ausencia de sesgo) ≥ 0.80

Resultados de la validación del contenido

Imagen	**IVC_Rep**	**IVC_Cal**	**IVC_Ses**
Dermatitis de contacto alérgica	0.8	1.0	0.8
Dermatitis atópica	1.0	1.0	0.8
Queratosis seborreica	1.0	1.0	1.0
Impétigo no bulloso	0.8	0.8	0.8
Psoriasis	1.0	1.0	0.8
Melanoma	0.6	1.0	0.8
Acné vulgar	1.0	0.8	1.0
Verruga vulgar	1.0	1.0	0.8
Rosácea	1.0	0.8	1.0
Tiña corporis	0.8	0.8	0.8
Carcinoma basocelular	1.0	0.8	1.0
Nevo melanocítico común	0.8	1.0	0.8
Lentigo solar	0.8	1.0	0.8
Queratosis actínica	1.0	1.0	1.0
Molusco contagioso	1.0	1.0	1.0

IVC: Índice de Validez de Contenido
Rep: Representatividad
Cal: Calidad técnica
Ses: Ausencia de sesgo

Fuente: Elaboración propia

El IVC global del instrumento se calculó como:

$$IVC_{Global} = \frac{\sum IVC_i}{k}$$

Donde:

- IVC_i = IVC de cada ítem
- k = número total de ítems (k = 15)

Resultados de la validación de expertos

IVC promedio por criterio	
Representatividad diagnóstica	0.907
Calidad técnica fotográfica	0.933
Ausencia de sesgo visual	0.880
IVC GLOBAL	0.907

Fuente: Elaboración propia

Interpretación de resultados

- **Validez general del instrumento:**

El IVC global de 0.907 indica una excelente validez de contenido para el instrumento en su conjunto (>0.90). De las 15 imágenes evaluadas, 14 (93.3%) cumplen simultáneamente con IVC ≥ 0.80 en los tres criterios. Solo una imagen (Melanoma, 6.7%) presenta deficiencia en un criterio (Representatividad diagnóstica).

Punto de desacuerdo: Lesiones pigmentadas

Durante la evaluación inicial del juicio de expertos, tres de los cinco expertos dieron puntajes poco favorables a las imágenes de nevus melanocítico común y melanoma, por lo que se decide cambiarlas por imágenes más representativas, de mejor calidad y con menor riesgo de sesgo, dos de los tres expertos aprobaron dicho cambio y las evaluaron con puntuaciones más favorables.

Pese a dicho cambio, uno de los expertos rechazó las imágenes de melanoma y nevus melanocítico común, dándoles una puntuación de 2 y 3 puntos respectivamente. Este desacuerdo genera la única deficiencia de validez en el instrumento.

El experto sugirió que un médico de atención primaria no debe intentar diferenciar melanoma de nevus benigno, sino que su responsabilidad clínica es reconocer lesiones melanocíticas sospechosas y derivar estos casos al especialista. En sus observaciones cuestiona la propiedad pedagógica de incluir diferencial melanoma/nevus en instrumento para estudiantes.

La imagen de Melanoma presenta desacuerdo legítimo en el criterio de representatividad diagnóstica (IVC=0.60), reflejando diferencias epistemológicas entre expertos sobre qué debe enseñarse a médicos generales. No obstante:

- Los criterios de calidad técnica y ausencia de sesgo son aceptables
- Hay consenso mayoritario de expertos
- La relevancia clínica es indiscutible

Por lo tanto se decide mantener la imagen y continuar con la evaluación de dichas patologías.

Conclusión sobre validez de contenido

El instrumento de evaluación diagnóstica basado en 15 imágenes clínicas fotográficas demuestra validez de contenido excelente conforme a estándares internacionales.

Fortalezas:

- IVC global de 0.907 (muy superior al umbral de 0.80)
- 14/15 imágenes (93.3%) con validez integral en los tres criterios
- Consenso perfecto en aspectos técnicos (IVC=0.933)
- Panel diverso y altamente cualificado (5 dermatólogos certificados, 5-42 años experiencia)

Limitaciones y consideraciones:

- Una imagen (6.7%) genera desacuerdo importante en un criterio

- Desacuerdo tiene fundamento epistemológico legítimo
- Requiere monitoreo durante implementación

Coeficiente de concordancia de Kendall

Adicional al Índice de Validez de Contenido (IVC), se empleó el coeficiente de concordancia de Kendall (W) como medida complementaria de acuerdo general entre expertos. Mientras que el IVC evalúa la validez de cada ítem de forma individual, el coeficiente W de Kendall cuantifica el grado de concordancia entre los cinco expertos en su rendimiento global de las imágenes. Esto permite determinar si los expertos, en promedio, clasifican las imágenes en orden similar respecto a su calidad, independientemente de que cada imagen cumpla con el criterio individual de IVC ≥ 0.80.

La fórmula del coeficiente W de Kendall es:

$$W = \frac{12 \times S}{k^2 \times (n^3 - n)}$$

Donde:

- S es la suma de cuadrados de las desviaciones de la suma de rangos respecto a su media ($S = \Sigma(R_i - \bar{R})^2$)
- k es el número de expertos (k=5)
- n es el número de ítems evaluados (n=15). El rango promedio esperado se calcula como $\bar{R} = k(n+1)/2$.

La interpretación del coeficiente W de Kendall sigue los siguientes criterios: W > 0.75 indica concordancia excelente; W entre 0.60 y 0.75 indica concordancia buena; W entre 0.40 y 0.60 indica concordancia moderada; y W < 0.40 indica concordancia pobre. Se realizó prueba de significancia estadística mediante el estadístico $x^2 = k(n-1)W$ con n - 1 grados de libertad, considerando significativa una concordancia cuando p < 0.05.

Coeficiente W de Kendall

Criterio	W	χ^2	p-value	Interpretación
Representatividad diagnóstica	0.6431	45.02	0.00004	Concordancia buena
Calidad técnica fotográfica	0.5980	41.86	0.00013	Concordancia moderada
Ausencia de sesgo visual	0.6943	48.60	0.00001	Concordancia buena

Fuente: Elaboración propia

Los coeficientes de concordancia de Kendall obtenidos fueron: W = 0.6431 ($p < 0.0001$) para representatividad diagnóstica, W = 0.5980 ($p < 0.0001$) para calidad técnica fotográfica, y W = 0.6943 ($p < 0.0001$) para ausencia de sesgo visual. Todos los valores de p fueron altamente significativos ($p < 0.001$), indicando que la concordancia observada entre expertos no es debida al azar.

En criterios de representatividad diagnóstica (W = 0.6431) y ausencia de sesgo visual (W = 0.6943), se alcanzó concordancia buena según los estándares de interpretación ($W > 0.60$). En calidad técnica fotográfica, la concordancia fue moderada (W = 0.5980, cercana al umbral de 0.60).

Esta variabilidad en concordancia refleja que, mientras los expertos muestran acuerdo general sobre la validez de las imágenes usadas en la encuesta, existe mayor consistencia en criterios técnicos y de ausencia de sesgo que en criterios más subjetivos como representatividad diagnóstica. Sin embargo, el rango de valores W = 0.60-0.69 y la significancia estadística universal (todos $p < 0.001$) confirman que los expertos, como grupo, convergen en sus evaluaciones de forma no aleatoria y estadísticamente robusta.

Ambas medidas de validez (IVC por ítem individual y W de Kendall por concordancia global) convergen en su conclusión: el instrumento de 15 imágenes clínicas fotográficas posee validez de contenido sólida, avalada por consenso sustancial entre expertos (IVC ≥ 0.80) y concordancia estadísticamente significativa.

7.7. Anexo 7: Validación psicométrica

Muestra del piloto

Se realizó un estudio piloto con 30 egresados de medicina de la Universidad Nacional Mayor de San Marcos. La muestra piloto incluyó participantes con las siguientes características:

- Edad promedio: 27.0 años (DE = 2.3, rango 23-31).
- Distribución por sexo: 66.7% femenino, 33.3% masculino.
- Distribución por institución de rotación: MINSA (44.4%), EsSalud (33.3%), instituciones privadas (22.2%).

Análisis de dificultad de ítems

El análisis de dificultad evalúa la proporción de participantes que respondieron correctamente cada ítem, permitiendo identificar preguntas excesivamente fáciles o difíciles. El índice de dificultad se calculó como:

$$p_i = \frac{Número\ de\ respuestas\ correctas\ en\ el\ ítem\ i}{Total\ de\ participantes}$$

Interpretación del índice de dificultad:

- p < 0.20: Ítem muy difícil (posiblemente mal diseñado)
- 0.20 ≤ p < 0.40: Ítem difícil
- 0.40 ≤ p < 0.60: Ítem de dificultad moderada (óptimo)
- 0.60 ≤ p < 0.80: Ítem fácil
- p ≥ 0.80: Ítem muy fácil (posiblemente trivial)

Rango ideal: 0.30 ≤ p ≤ 0.80

Resultados del análisis de dificultad:

Nivel de dificultad por ítems

Ítem	Diagnóstico	Correctas	Total	p	Categoría
P1	Dermatitis de contacto alérgica	21	30	0.700	Fácil
P2	Dermatitis atópica	18	30	0.600	Fácil
P3	Queratosis seborreica	14	30	0.467	Moderada
P4	Impétigo no bulloso	25	30	0.833	Muy fácil
P5	Psoriasis	23	30	0.767	Fácil
P6	Melanoma	21	30	0.700	Fácil
P7	Acné vulgar	24	30	0.800	Muy fácil
P8	Verruga vulgar	19	30	0.633	Fácil
P9	Rosácea	25	30	0.833	Muy fácil
P10	Tiña corporis	21	30	0.700	Fácil
P11	Carcinoma basocelular	23	30	0.767	Fácil
P12	Nevo melanocítico común	20	30	0.667	Fácil
P13	Lentigo solar	22	30	0.733	Fácil
P14	Queratosis actínica	17	30	0.567	Moderada
P15	Molusco contagioso	23	30	0.767	Fácil

Fuente: Elaboración propia

El índice de dificultad promedio fue p = 0.702, indicando un balance apropiado entre ítems fáciles y difíciles. Del total de 15 ítems, 13 (86.7%) se encuentran en el rango ideal (0.30-0.80), dos ítems (P4 y P9) mostraron facilidad elevada (p > 0.80) pero aún dentro de parámetros aceptables, y ningún ítem resultó excesivamente difícil (p<0.20).

Análisis de Discriminación de Ítems

El análisis de discriminación evalúa la capacidad de cada ítem para diferenciar entre participantes con alto y bajo rendimiento general. Se utilizó la correlación punto-biserial (r_{pbis}), que mide la correlación entre el desempeño en un ítem específico (variable dicotómica) y el puntaje total del test (variable continua). La fórmula es:

$$r_{pbis} = \frac{(M_1 - M_0)\sqrt{p(p-1)}}{SD_{Total}}$$

Donde:

- M_1 = Media del puntaje total de quienes respondieron correctamente el ítem
- M_0 = Media del puntaje total de quienes respondieron incorrectamente el ítem
- p = Proporción que respondió correctamente el ítem
- SD_{Total} = Desviación estándar de los puntajes totales

Interpretación del r_{pbis}:

- r_{pbis} > 0.30: Buena discriminación (ítem funciona adecuadamente)
- 0.20 ≤ r_{pbis} ≤ 0.30: Discriminación aceptable (considerar revisión)
- r_{pbis} < 0.20: Discriminación pobre (revisar o eliminar)
- r_{pbis} < 0: Discriminación negativa (posible error en clave de respuesta)

Resultados del análisis de discriminación:

Estadísticas globales:

- Media de aciertos: 10.53 de 15 (70.2%)
- Desviación estándar: 2.99
- Rango: 0-15 aciertos

Análisis de discriminación de ítems

Ítem	M_1 (correctos)	M_0 (incorrectos)	rpbis	Categoría
P1	11.43	8.44	0.457	Buena
P2	11.83	8.58	0.532	Buena
P3	12.36	8.94	0.570	Buena
P4	11.08	7.80	0.409	Buena
P5	11.13	8.57	0.362	Buena
P6	11.43	8.44	0.457	Buena

P7	11.00	8.67	0.312	Buena
P8	11.16	9.45	0.274	Aceptable
P9	11.16	7.40	0.468	Buena
P10	11.52	8.22	0.506	Buena
P11	11.35	7.86	0.494	Buena
P12	11.15	9.30	0.292	Aceptable
P13	11.41	8.12	0.486	Buena
P14	11.76	8.92	0.471	Buena
P15	11.17	8.43	0.388	Buena

Fuente: Elaboración propia

La correlación punto-biserial promedio fue r_{pbis} = 0.432, indicando excelente poder discriminativo general. Del total de 15 ítems, 13 (86.7%) presentaron buena discriminación (r_{pbis} > 0.30) y 2 (13.3%) discriminación aceptable (0.20 ≤ r_{pbis} ≤ 0.30). No se identificaron ítems con discriminación pobre o negativa, confirmando que todos los ítems funcionan adecuadamente para diferenciar entre participantes con diferentes niveles de precisión diagnóstica.

Confiabilidad Interna (Coeficiente KR-20)

Para evaluar la confiabilidad interna del instrumento, se calculó el coeficiente Kuder-Richardson 20 (KR-20), específicamente diseñado para tests con respuestas dicotómicas (correcto/incorrecto). El KR-20 mide la consistencia interna, es decir, el grado en que los ítems del test se correlacionan entre sí y miden un constructo común. La fórmula es:

$$KR20 = \frac{k}{k-1} \times \left[1 - \frac{\sum p_i q_i}{\delta^2}\right]$$

Donde:

- k = Número de ítems (k = 15)
- p_i = Proporción de respuestas correctas en el ítem *i*
- $q_i = 1 - p_i$ = Proporción de respuestas correctas en el ítem *i*

- $q_i p_i$ = Varianza del ítem *i*
- $\sum p_i q_i$ = Suma de varianzas de todos los ítems
- δ^2 = varianza total de los puntajes del test

Procedimiento de cálculo paso a paso:

Paso 1: Calcular p_i, q_i y $q_i p_i$ para cada ítem.

Ítem	Correctas	Total	p_i	q_i	$q_i p_i$
P1	21	30	0.700	0.300	0.2100
P2	18	30	0.600	0.400	0.2400
P3	14	30	0.467	0.533	0.2489
P4	25	30	0.833	0.167	0.1389
P5	23	30	0.767	0.233	0.1789
P6	21	30	0.700	0.300	0.2100
P7	24	30	0.800	0.200	0.1600
P8	19	30	0.633	0.367	0.2322
P9	25	30	0.833	0.167	0.1389
P10	21	30	0.700	0.300	0.2100
P11	23	30	0.767	0.233	0.1789
P12	20	30	0.667	0.333	0.2222
P13	22	30	0.733	0.267	0.1956
P14	17	30	0.567	0.433	0.2456
P15	23	30	0.767	0.233	0.1789

Fuente: Elaboración propia

Paso 2: Sumar todas las varianzas individuales.

$$\sum p_i q_i = 2.9889$$

Paso 3: Calcular la varianza total de los puntajes.

Media de puntajes totales: $\overline{X}$ =10.53 (de 15 posibles)

$$\delta^2 = \frac{\sum(X_i - X)^2}{n-1} = 8.9471$$

Paso 4: Aplicar la fórmula de KR-20.

$$KR20 = \frac{15}{15-1} \times \left[1 - \frac{2.9889}{8.9471}\right]$$

$$KR20 = 1.0714 \times [1 - 0.3341]$$

$$KR20 = 1.0714 \times 0.6659 = 0.7135$$

Interpretación del coeficiente KR-20:

- KR-20 < 0.50: Confiabilidad inaceptable
- 0.50 ≤ KR-20 < 0.60: Confiabilidad marginal
- 0.60 ≤ KR-20 < 0.70: Confiabilidad aceptable
- 0.70 ≤ KR-20 < 0.80: Confiabilidad excelente
- KR-20 ≥ 0.80: Confiabilidad muy alta (posiblemente excesivamente homogénea)

Resultado: El coeficiente KR-20 obtenido fue 0.7135, ubicándose en la categoría de confiabilidad excelente. Este valor indica que el 71.35% de la varianza en los puntajes es atribuible a diferencias reales en precisión diagnóstica, mientras que solo el 28.65% corresponde a varianza de error. El instrumento demuestra consistencia interna apropiada para medir la precisión diagnóstica en evaluación dermatológica.

Método de Validación	Estadístico	Valor Obtenido	Interpretación
Validez de Contenido	IVC global	0.907	Excelente
Concordancia entre expertos	W de Kendall	0.64-0.69	Buena
Dificultad de ítems	*p* promedio	0.702	Balanceada
Ítems en rango ideal	Proporción	13/15 (86.7%)	Excelente
Discriminación de ítems	*rpbis* promedio	0.432	Excelente

Ítems con buena discriminación	Proporción	13/15 (86.7%)	Excelente
Confiabilidad interna	KR-20	0.7135	Excelente

Fuente: Elaboración propia

Conclusiones de la validación:

- **Validez de contenido:** El instrumento demostró validez de contenido excelente (IVC = 0.907), con 93.3% de las imágenes validadas por consenso de expertos. El coeficiente de concordancia de Kendall confirmó un acuerdo estadísticamente significativo entre los cinco evaluadores.
- **Validez de constructo:** El análisis de dificultad reveló distribución apropiada de ítems, con índice promedio de 0.702 y 86.7% de ítems en rango ideal (0.30-0.80), confirmando que el instrumento presenta desafío diagnóstico apropiado sin preguntas triviales o excesivamente difíciles.
- **Poder discriminativo:** La correlación punto-biserial promedio de 0.432 indica que los ítems diferencian adecuadamente entre participantes con alta y baja precisión diagnóstica. El 86.7% de ítems presentaron discriminación buena (r_{pbis} > 0.30), sin ítems problemáticos.
- **Confiabilidad:** El coeficiente KR-20 de 0.7135 confirma excelente consistencia interna, indicando que el instrumento mide de forma coherente y reproducible la precisión diagnóstica dermatológica.

La convergencia de todos los indicadores psicométricos valida el instrumento como herramienta confiable y válida para evaluar la precisión diagnóstica en dermatología mediante imágenes clínicas fotográficas. El instrumento está listo para su implementación en la fase de recolección de datos finales.

7.8. Anexo 8: Acta de aprobación del proyecto por el Comité de Ética de Investigación

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
(Universidad del Perú DECANA DE AMÉRICA)
FACULTAD DE MEDICINA
INSTITUTO DE ÉTICA EN SALUD
COMITÉ DE ÉTICA DE INVESTIGACIÓN
"Año del Fortalecimiento de la Soberanía Nacional"

ACTA DE EVALUACIÓN ÉTICA DE ESTUDIOS DE INVESTIGACIÓN
(AEE-CEI)

CÓDIGO DE ESTUDIO N°: 0102-2026

En Lima, a los diecisiete días del mes de febrero, en Sesión del COMITÉ DE ÉTICA EN INVESTIGACIÓN, previa evaluación del Proyecto de Tesis de Pregrado titulado: **"Precisión diagnóstica en dermatología mediante imágenes clínicas en egresados de medicina de la UNMSM, 2025"** presentado por **Walter Allccarima Ramos** con código 18010285, de la escuela profesional de medicina humana, para optar el Título Profesional de Médico Cirujano en la Facultad de Medicina Humana de la Universidad Nacional Mayor de San Marcos.

ACUERDA:

Dar por **APROBADO** dicho Proyecto, considerando que se ha cumplido satisfactoriamente con las recomendaciones en aspectos Científicos Técnicos y Éticos para la investigación en seres humanos.

"El presente documento tiene vigencia a partir de la fecha y expira el 16 de febrero de 2027"

Lima, 17 de febrero de 2026

JUAN CARLOS OCAMPO ZEGARRA
PRESIDENTE DEL C.E.I. / F.M. / U.N.M.S.M.
FACULTAD DE MEDICINA SAN FERNANDO
CODIGO DOCENTE: 0A3079
CMP: 043040 - RNE: 028980 - RNSE: S00415

Dr. Juan Carlos Ocampo Zegarra
Presidente del CEI/FM/UNMSM