

# UNIVERSITAT DE VALÈNCIA



## FACULTAT DE MEDICINA I ODONTOLOGIA

Departamento de Medicina

Programa de Doctorado R.D. 99/2011

Año 2020

### **TESIS DOCTORAL:**

RECURRENCIA POSTQUIRÚRGICA EN LA ENFERMEDAD  
DE CROHN. ANÁLISIS DE LA VARIABILIDAD INTRA E  
INTEROBSERVADOR DEL ÍNDICE DE RUTGEERTS.

Presentada por: D<sup>a</sup> Noelia Alonso Lázaro

Directores:

Dr. Vicente Pons Beltrán.

Dra. Pilar Nos Mateu.

Tutor: Dr. Vicente Garrigues Gil.



D/ **VICENTE PONS BELTRÁN**, Jefe de Sección de Endoscopias Digestivas del Servicio de Medicina Digestiva del Hospital Universitari i Politècnic La Fe de Valencia.

Dña **PILAR NOS MATEU**, Jefa del Servicio de Medicina Digestiva del Hospital Universitari i Politècnic La Fe de Valencia.

D/**VICENTE GARRIGUES GIL**, Jefe de Sección. Servicio de Medicina Digestiva. Hospital Universitari i Politècnic La Fe. Valencia. Profesor Titular. Departament de Medicina. Facultat de Medicina i Odontologia. Universitat de València.

**CERTIFICAN:**

Que la presente tesis, titulada “**RECURRENCIA POSTQUIRÚRGICA EN LA ENFERMEDAD DE CROHN. ANÁLISIS DE LA VARIABILIDAD INTRA E INTEROBSERVADOR DEL ÍNDICE DE RUTGEERTS**”, corresponde al trabajo realizado bajo su dirección por Dña. **NOELIA ALONSO LÁZARO**, y cumple los méritos suficientes para ser evaluada por el tribunal correspondiente para poder optar al grado de Doctor por la Universitat de València.

Y para que conste firman el presente certificado en Valencia, a 11 de JUNIO de 2020.

Fdo. Dr. VICENTE  
PONS BELTRÁN

Fdo. Dra PILAR  
NOS MATEU

Fdo. Dr. VICENTE  
GARRIGUES GIL



A mi madre, esta tesis y parte de mi carrera profesional, sin duda es tuya. Gracias por dármele todo sin pedir nada a cambio.

A mi padre, por inculcarme desde pequeña el espíritu de trabajo y valorar, aunque fuera de golpe, que sin la familia no somos nada.

A mi hermana, mi segunda madre, por su fortaleza y su apoyo incondicional a pesar de sus circunstancias.

A Vicent, mi mejor amigo y mi compañero de viaje, por su cariño, su paciencia y su serenidad, por ser mi complemento perfecto.

Y por supuesto a Sara, mi tesoro, porque esas risas han sido el motor que me ha impulsado a acabar este proyecto.

## AGRADECIMIENTOS

A Vicente Pons, co-director de esta tesis y responsable de la Unidad de Endoscopia Digestiva. Mi eterna gratitud por la confianza depositada en mí cuando ni yo misma la tenía. Gracias por enseñarme todo lo que sé y por convertir a la endoscopia en parte de mi mundo. Nunca podré agradecerle lo suficiente todas las oportunidades que me has brindado.

A Pilar Nos, co-directora de esta tesis y Jefa del servicio de Medicina Digestiva. Gracias por adentrarme en el mundo de la enfermedad inflamatoria, por transmitirme esa pasión y ese espíritu de trabajo y por dedicarme el tiempo aun sacándolo debajo de las piedras. Soy consciente del esfuerzo y lo valoro y agradezco enormemente.

Sin vuestro apoyo incondicional ese trabajo hubiera sido imposible de realizar.

A Vicente Garrigues, tutor de esta tesis, muchas gracias por tu disponibilidad, tus comentarios y correcciones, siempre tan acertadas.

A la Unidad de Enfermedad Inflamatoria Intestinal por su participación de forma activa en esta tesis, por haber dedicado parte de vuestro tiempo libre en este proyecto, aun suponiendo una carga importante, habéis sido un pilar fundamental.

A mis compañeros y amigos de endoscopias, mi eterna gratitud no sólo por la participación en la tesis sino por el apoyo constante y por todo el aprendizaje que me habéis dado. Me siento afortunada de pertenecer a esta pequeña familia y seguir aprendiendo y creciendo como endoscopista a vuestro lado. En especial a Marco Bustamante, por esos “cables de última hora”, él ya sabe a lo que me refiero.

A mis amigas y compañeras de viaje, Caren, Carla y Andrea, por ese apoyo constante y por ser los pilares en los que he podido apoyarme en muchos momentos duros.

A los residentes que participaron en el proyecto, convertidos en grandes adjuntos y grandes amigos. Gracias no sólo por la disponibilidad sin horarios y el afán de aprender sino por ser el estímulo que me ha impulsado a mejorar día a día, sin vosotros hubiera sido imposible.

## COLABORACIONES

Pau Miro i Martinez y Jorge Jordan Nuñez, Profesores del Departamento de Estadística e Investigación Operativa Aplicadas y Calidad de la Universitat Politècnica de València, que han colaborado intensamente en el diseño del trabajo y en el análisis estadístico de este proyecto. Este proyecto no hubiera sido posible sin su desinteresada participación.





## ÍNDICE

AGRADECIMIENTOS .....	6
COLABORACIONES .....	7
LISTA DE ABREVIATURAS .....	11
FIGURAS.....	13
TABLAS .....	15
1. INTRODUCCIÓN .....	17
2. REVISIÓN BIBLIOGRÁFICA .....	21
2.1. Enfermedad Inflamatoria Intestinal.....	23
2.1.1 Definición.....	23
2.1.2. Epidemiología. ....	23
2.1.3. Fisiopatología y etiopatogenia.....	25
2.1.4. Manifestaciones clínicas. ....	29
2.1.5. Criterios diagnósticos.....	31
2.1.6. Clasificación.....	34
2.1.7. Valoración de la actividad inflamatoria .....	36
2.2 Cirugía en la Enfermedad de Crohn: Recurrencia.....	43
2.2.1. Introducción .....	43
2.2.2. Recurrencia: definiciones .....	44
2.2.3 Historia natural: Recurrencia post-quirúrgica. ....	45
2.2.4. Valoración de la recurrencia: métodos disponibles.....	47
2.2.5. Prevención de la Recurrencia .....	50
2.2.6. Seguimiento de la EC.....	55
2.3. Endoscopia en la EC .....	58
2.3.1. Descripción de los índices endoscópicos .....	58
2.3.2. Curación mucosa.....	65
2.4. Valoración de los Índices de Actividad. Función del Lector .....	69
2.5. Proceso de aprendizaje en lectura endoscópica.....	76
2.5.1 Definición de la evaluación de competencias en la endoscopia digestiva. ....	76
2.5.2 Evaluación de habilidades cognitivas, valoración de lesiones. Curva de aprendizaje con material visual. ....	83
2.5.3. Competencias en endoscopia en la EII.....	85
3. JUSTIFICACIÓN .....	87

4. HIPÓTESIS.....	91
5. OBJETIVOS .....	95
6. MATERIAL Y MÉTODO .....	99
6.1. Diseño del estudio .....	101
6.2. Material y Métodos. ....	101
6.2.1 Equipamiento endoscópico .....	101
6.2.2. Selección de los vídeos: criterios de inclusión y exclusión.....	103
6.2.3. Plataforma de almacenamiento: Plataforma Owncloud la fe .....	106
6.2.4. Plataforma para la recogida de datos: Formularios Google. ....	106
6.2.5. Lectores .....	108
6.2.6. Formación.....	110
6.2.7. Variables del estudio. ....	110
6.2.8. Evolución cronológica del estudio. ....	111
7. ANÁLISIS ESTADÍSTICO .....	113
8. RESULTADOS.....	119
9. DISCUSIÓN .....	151
10. LIMITACIONES .....	175
11. CONCLUSIONES .....	179
12. ANEXOS.....	183
Anexo 1: Documento de aprobación.....	185
Anexo 2: Formulario de lectura de los vídeos .....	186
Anexo 3: Hoja basal de recogida de datos. ....	192
Anexo 4: Variabilidad intraobservador .....	194
Anexo 5: Resultados con un experto en EII como lector central. ....	197
13. BIBLIOGRAFÍA.....	199

## LISTA DE ABREVIATURAS

ADA: Adalimumab.

Anti-TNF: anti-Tumoral Necrosis Factor.

CCI coeficiente de correlación intraclase.

CDAI: Crohn's disease activity index.

CDEIS: Crohn disease endoscopic index activity.

CF: calprotectina fecal.

CU: Colitis ulcerosa.

EC: Enfermedad de Crohn.

ECCO: European Crohn's and Colitis Organisation.

EII: Enfermedad inflamatoria intestinal.

Entero-RM: enterografía por resonancia magnética.

Entero-TAC: enterografía por tomografía axial computarizada.

EpiCom: Epidemiological Committee.

EXTEND: Extend the Safety and Efficacy of Adalimumab through Endoscopic Healing.

GETECCU: Grupo Español de trabajo en Enfermedad de Crohn y Colitis Ulcerosa.

FDA: Food and Drug Administration.

IFX: Infliximab.

MIR: médico interno residente en formación.

PCR: proteína C reactiva.

POCER: Postoperative Crohn Endoscopic Recurrence.

RPQ: recurrencia postquirúrgica.

SES-CD: Simple endoscopic score Crohn Disease.



## FIGURAS

Figura 1. Probabilidad de intervención quirúrgica en EC dependiendo del año de diagnóstico y tratamientos disponibles.

Figura 2: Algoritmo de seguimiento de pacientes con EC intervenida según la American Gastroenterological Association.

Figura 3. Algoritmo de seguimiento de EC intervenida según GETECCU.

Figura 4. Esquema de la adquisición de competencias en la Endoscopia Digestiva.

Figura 5. Diferencias entre sistema tradicional de aprendizaje y el basado en competencias.

Figura 6. Fases del aprendizaje.

Figura 7. Sala de trabajo de la Unidad de Endoscopia Digestiva.

Figura 8. Procesador EXERA III. Procesador EXERA II.

Figura 9. Endoscopio Olympus CF-180AL.

Figura 10. Tipos de anastomosis ileocólicas

Figura 11. Plataforma owncloud

Figura 12.1. Primera parte del Formulario 1. 12.2. Segunda parte del Formulario 1.

Figura 13. Página de formulario.

Figura 14. Diagrama cronológico del estudio.

Figura 15. Resumen del análisis de concordancia realizado entre los lectores.

Figura 16. Puntuación de los vídeos según el índice de Rutgeerts modificado calculado por todos los lectores.

Figura 17. Concordancia según el valor kappa de los lectores comparado con el lector central.

Figura 18. Porcentaje de acuerdo de los lectores expertos comparado con los lectores centrales.

Figura 19 Porcentaje de acuerdo de los lectores comparado con los diferentes lectores centrales.

Figura 20. Distribución de la puntuación del índice de Rutgeerts modificado de los lectores no expertos pre-formación.

Figura 21. Distribución de la puntuación del índice de Rutgeerts modificado de los lectores no expertos post-formación.

Figura 22. Representación gráfica de los valores kappa pre y post-formación de forma global.

Figura 23. Representación gráfica del porcentaje de acuerdos pre y postformación.

Figura 24. Representación gráfica de los valores kappa pre y post-formación comparando los lectores con el lector central.

Figura 25. Representación gráfica del porcentaje de acuerdo pre y post-formación comparando los lectores con el lector central.

Figura 26. Porcentaje de lectores que difirieron en las puntuaciones de cada categoría del índice de Rutgeerts en comparación con el lector central.

## TABLAS

Tabla 1. Criterios de Lennard-Jones.

Tabla 2. Clasificación de Viena EC.

Tabla 3. Clasificación de Montreal.

Tabla 4. Índice de Sailer.

Tabla 5. Riesgo de recurrencia clínica y endoscópica de la EC sin tratamiento en función de los factores de riesgo.

Tabla 6. Descripción gráfica del CDEIS.

Tabla 7. SES-CD.

Tabla 8. Índice endoscópico de Rutgeerts. Índice de Rutgeerts modificado.

Tabla 9. Resumen de algunas definiciones de respuesta endoscópica, remisión endoscópica y curación mucosa.

Tabla 10. Distribución inicial de los vídeos según el índice de Rutgeerts modificado otorgado por el lector local.

Tabla 11. Características demográficas de los pacientes incluidos.

Tabla 12. Características quirúrgicas de los pacientes incluidos.

Tabla 13. Concordancia interobservador por grupos de expertos y globalmente.

Tabla 14. Variabilidad global y por expertos según las agrupaciones del índice de Rutgeerts modificado.

Tabla 15. Variabilidad comparada con el lector central del grupo de expertos.

Tabla 16. Variabilidad del grupo de expertos comparada con el lector central.

Tabla 17. Porcentaje de acuerdo de los lectores expertos comparado con el lector central.

Tabla 18. Concordancia del índice de Rutgeerts agruado en el grupo de expertos global y por categorías.

Tabla 19. Concordancia según varias agrupaciones del índice de Rutgeerts modificado en el grupo de expertos global, comparado con el grupo de expertos con el lector central.

Tabla 20. Variabilidad interobservador pre y post-formacion determinada por el índice kappa.

Tabla 21. Variabilidad interobservador de la lectura post-formación, comparada con el lector central.

Tabla 22. Variabilidad interobservador del índice de Rutgeerts modificado agrupado, pre y post-formación.

Tabla 23. Variabilidad interobservador del índice de Rutgeerts modificado agrupado pre y postformación comparado con el lector central.

Tabla 24. Diferencias entre la lectura pre-formación de lectores no expertos comparada con el grupo de expertos globalmente en las distintas agrupaciones.

Tabla 25. Diferencias entre la lectura post-formación de lectores no expertos comparada con los expertos.

Tabla 26. Esquemas de lectura.



## 1. INTRODUCCIÓN

---



La enfermedad de Crohn (EC) es una enfermedad inflamatoria intestinal crónica de etiología desconocida que cursa con brotes afectando de forma muy importante la calidad de vida de los pacientes. Durante el desarrollo de la enfermedad, la mayoría de los pacientes van a precisar al menos una intervención quirúrgica; sin embargo, la cirugía no es curativa y la reaparición de las lesiones en la propia anastomosis o bien en el neoíleon (vertiente ileal de la anastomosis) sin afectación previa es la norma. Por ello, son necesarias técnicas para valorar la recurrencia de la enfermedad con el objetivo de instaurar y monitorizar el tratamiento más adecuado en función de la gravedad.

El método de elección para valorar la recurrencia endoscópica es la colonoscopia. A pesar de que anteriormente se utilizaban los índices de actividad clínicos para incluir a los pacientes y valorar la eficacia de los fármacos en los ensayos clínicos, actualmente con el objetivo de alcanzar la curación mucosa, se están incorporando los índices de actividad endoscópicos. Entre ellos, el índice de Rutgeerts es el utilizado no sólo en los ensayos clínicos sino en la práctica clínica diaria para valorar la recurrencia de la EC. Se recomienda realizar una endoscopia a los 6 meses de la intervención ya que se ha observado que lesiones graves en la endoscopia predicen un curso más agresivo de la enfermedad, teniendo implicaciones terapéuticas importantes. Según el índice de Rutgeerts, en la mayoría de los estudios se considera recurrencia un índice superior a i2; sin embargo, los pacientes que presentan este score tienen un curso de la enfermedad menos predecible que los que presentan un score i3 o i4 cuyo pronóstico es claramente peor. En el score i2 se incluyen las lesiones confinadas a la anastomosis. En algunos trabajos se ha sugerido que estas lesiones podrían ser debido a fenómenos isquémicos relacionados con la propia técnica quirúrgica sin una traducción clínica relevante. Por ello, posteriormente se realizó una subclasificación de este índice, el índice de Rutgeerts modificado. diferenciando i2a que englobaría las lesiones localizadas en la anastomosis con menos probabilidad de traducir un impacto clínico, e i2b, que reflejaría la presencia de más de 5 aftas en el neoíleon con mucosa normal entre ellas, o una afectación salteada (áreas no contiguas con lesiones más intensas). Todos los índices, incluyendo el índice de Rutgeerts, se basan en una serie de variables que en muchas ocasiones tienen un componente subjetivo explorador dependiente, por lo que la puntuación obtenida va a depender del endoscopista que realiza la colonoscopia. En los trabajos publicados hasta ahora queda reflejada la amplia

variabilidad interobservador de este índice y por tanto, su baja reproducibilidad. Debido a ello, apareció la figura del lector central o el *off site reader*, una figura ajena a la información clínica de los pacientes que realiza la lectura de las colonoscopias y asigna un índice de actividad sólo con las imágenes endoscópicas sin disponer de más datos. Se supone, por razones teóricas, que va a aportar más objetividad a los estudios reduciendo la variabilidad interobservador que existe en la utilización de los propios índices. A pesar de que en otras especialidades se ha visto su utilidad, no se dispone de suficientes datos para recomendar su implementación en la enfermedad inflamatoria intestinal y concretamente en la EC.

Este estudio pretende profundizar en los datos a favor y en contra de la utilización generalizada del índice de Rutgeerts modificado en los ensayos clínicos controlados y en la práctica clínica, abordando la importancia y discrepancias de su lectura tanto centralizada como por grupos diferenciados (endoscopistas, expertos en enfermedad inflamatoria intestinal y médicos en formación); así como las ventajas de programas docentes de entrenamiento. El estudio presentado analiza los datos de los pacientes con EC intervenida en los que se ha realizado una anastomosis ileocólica y que disponen de una colonoscopia adecuada para valorar la recurrencia post-quirúrgica utilizando el índice de Rutgeerts modificado. Se incluyen a los pacientes en seguimiento en consultas externas de Gastroenterología del Hospital Universitari i Politècnic La Fe a los que se les solicita la colonoscopia como seguimiento habitual de su enfermedad. Se pretende analizar la variabilidad intraobservador e interobservador del índice de Rutgeerts modificado entre diferentes lectores, incluyendo el lector central, y valorar si la formación de los lectores no expertos mejora la concordancia del índice.

## 2. REVISIÓN BIBLIOGRÁFICA

---



## **2.1. Enfermedad Inflamatoria Intestinal.**

### **2.1.1 Definición.**

Dentro de la enfermedad inflamatoria intestinal (EII) se engloban tres enfermedades digestivas, la enfermedad de Crohn (EC), la colitis ulcerosa (CU) y la colitis indeterminada. Todas ellas son un grupo de enfermedades crónicas que afectan al tracto digestivo ocasionando un proceso inflamatorio crónico. Este proceso cursa con brotes de actividad, de gravedad variable, intercalados con periodos de remisión, siendo impredecible el número y la gravedad de los mismos.

A pesar de englobarlas dentro del mismo concepto, son entidades clínicas diferentes. La CU, afecta exclusivamente al colon, se limita a la mucosa y se inicia en el recto (casi siempre está afecto) alcanzando una extensión variable pero en continuidad.

Mientras que por otra parte está la EC, es una inflamación crónica transmural que puede afectar a cualquier tramo del tracto digestivo desde la boca hasta el ano, y que puede presentarse en forma de tres patrones, inflamatorio, estenosante y penetrante, lo que explica la formación de estenosis y fístulas<sup>1</sup>.

### **2.1.2. Epidemiología.**

La EII afecta a 5 millones de personas a nivel mundial, con mayor índice en Europa en comparación con Norteamérica, 3 vs 1,5 millones respectivamente<sup>2</sup>.

A lo largo del siglo XX se ha producido un incremento exponencial en la incidencia de la EII tanto en población pediátrica como en población adulta, sobre todo en los países industrializados y en los individuos de raza blanca<sup>3,4</sup>. Esto ha convertido a la EII en un problema de salud pública importante tanto por los costes directos como indirectos que conlleva. La incidencia más alta se encuentra en los países más desarrollados como los que se localizan en el Norte de Europa<sup>2,5</sup> y en Norteamérica. Actualmente se estima que la incidencia anual de EII en Europa es de 0,9-24/100.000 habitantes para CU y 11,5/100.000

habitantes para la EC. En Norteamérica los datos son similares, con una incidencia anual de 19,2/100.000 habitantes en el caso de la CU y de 20,2/100.000 habitantes en la EC.

La prevalencia se afecta tanto por la incidencia como por la duración de la enfermedad. La prevalencia de CU es muy similar en Europa y Norteamérica con cifras de 294 vs 249 por cada 100.000 habitantes respectivamente y en el caso de la EC existe mayor discrepancia, con una prevalencia de 213 vs 318 por cada 100.000 habitantes respectivamente.

Sin embargo, a lo largo del siglo XXI se ha observado un cambio en la tendencia. En algunas regiones parece que tiende a estabilizarse o incluso a descender<sup>6</sup>, mientras que en otras regiones, donde anteriormente no era frecuente, como Europa del Este, África, Asia y otras regiones del pacífico, se observa un incremento de la incidencia. Esto se puede atribuir al desarrollo tecnológico, a la mejora en la calidad de vida y por supuesto, al interés creciente por parte de los clínicos por el estudio de estas enfermedades<sup>2,7,8</sup>.

En Europa, recientemente se ha publicado un estudio elaborado por el *Epidemiological Committee* (EpiCom) en el que describen un gradiente Oeste-Este en cuanto a la incidencia de la EII. En este estudio, la incidencia combinada de la EC y la CU es de 5,4/100.000 habitantes y 8,2/100.000 habitantes respectivamente, siendo Faroe Islands el lugar con una incidencia más elevada (81,5/100.000 habitantes). De momento no disponemos de datos similares entre Europa y América, y además los que hay se basan en cohortes de determinadas poblaciones sin ser representativas de la población general<sup>5</sup>.

Sin embargo, en Asia sí que ha elaborado un estudio similar al EpiCom, el ACCESS (*Asia-Pacific Crohn's and Colitis Epidemiology Study*), encontrando una incidencia global anual de 0,8 y 0,5 /100.000 habitantes para CU y EC respectivamente en Asia, siendo China el país con más casos (3,44/100.000). En Australia la incidencia para la CU es de 7,3/100.000 habitantes y, algo superior para la EC<sup>7</sup>.

Recientemente se ha publicado una revisión sistemática sobre la incidencia global de la EII<sup>9</sup>. Todos los datos disponibles hasta la fecha coinciden en que a pesar de que la prevalencia de la enfermedad sigue siendo elevada en los países industrializados con cifras



que rondan el 0,3%, su incidencia se está estabilizando o incluso reduciendo en algunas regiones. Al contrario ocurre en regiones de África y Asia en los que prácticamente se desconocía la enfermedad, que con el desarrollo tecnológico y la globalización se está produciendo un incremento en la incidencia.

### 2.1.3. Fisiopatología y etiopatogenia.

El mecanismo etiopatogénico sigue siendo en su mayor parte desconocido. Existen diversas hipótesis, siendo la más aceptada la que se fundamenta en que existe una respuesta inflamatoria descontrolada y excesiva de la mucosa intestinal en individuos genéticamente predispuestos, debido a la interacción de un factor ambiental, “trigger”, con la microbiota intestinal.

Por tanto, según esta hipótesis, la etiopatogenia se centra en tres pilares fundamentales, la inmunidad, la genética y el ambiente. La EC es una enfermedad sistémica, cada factor por sí sólo no es capaz de producir la enfermedad, sin embargo, la interacción de todos ellos puede desencadenarla. Esta interacción es muy compleja, y se han realizado múltiples estudios, tanto genéticos como de microbiota intestinal para ampliar el nivel de conocimiento y realizar terapias más dirigidas<sup>10-14</sup>.

A pesar de desconocer la etiopatogenia concreta que pone en marcha la cascada inflamatoria que da lugar a la expresión de la EC, con el paso de los años y los estudios desarrollados en este campo, se va obteniendo cada vez más información que permite desarrollar terapias dirigidas a diferentes etapas de esa cascada inflamatoria.

1. Factores genéticos: la EC se considera una enfermedad poligénica. En estudios poblacionales se ha demostrado que el factor genético incrementa el riesgo de padecer EII<sup>15</sup>. La etnia judía Ashkenazi es la que mayor riesgo tiene en comparación con la población general<sup>16</sup> y estudios familiares han demostrado cómo los gemelos monocigotos tienen mayor grado de concordancia que los dicigotos y que existe agregación familiar, sobre todo, en familiares de primer grado<sup>11</sup>.

No obstante, la alteración en esta enfermedad es una alteración poligénica. Uno de los genes más estudiados y conocidos es el NOD 2-CARD15 (*Caspase-Recruitment-Domain-15*), que incrementa la susceptibilidad de padecer EC sobre todo en la etnia caucásica. Se encuentra localizado en el cromosoma 16 y su función está basada en la regulación de la respuesta inmune innata. Otros genes relacionados con la inmunidad innata se han relacionado con la EC, como el ATG16L1 (*autophagy-related 16-like-1*), IRGM (*Immunity-related GTPase family M protein*) o LRRK2 (*leucine rich repeat kinase 2*) que regulan la autofagia y también genes relacionados con vías de señalización de IL-23, como el gen que regula el receptor de la IL-23 o la subunidad p40 de la IL-12B<sup>17</sup>.

La mutación genética es importante, pero en muchas ocasiones su trascendencia clínica es escasa, ya que sólo explica un 13% de los casos de EC; siendo los factores epigenéticos y ambientales fundamentales para el desarrollo de la enfermedad.

2. Microbiota: se encuentra formada por una amplia población de microorganismos que habitan en la luz intestinal. Juega un papel fundamental en la digestión y sobre todo en la formación del sistema inmune entérico y en la protección de la homeostasis, por ser un componente fundamental de la barrera entérica. No obstante, con el cambio en el hábito alimentario, debido a la ingesta de gran cantidad de carbohidratos y de grasas, así como a la exposición a fármacos y otras drogas, la microbiota intestinal cambia produciéndose un desequilibrio entre organismos patógenos y protectores.

No se conoce con exactitud el mecanismo por el cual la microbiota participa en el proceso inflamatorio. Una de las hipótesis es la disbiosis, es decir, los cambios en la microbiota intestinal generarían una respuesta inmune patológica en pacientes susceptibles, y, por tanto, desarrollarían la inflamación intestinal.

Los pacientes con EII presentan un desequilibrio en los constituyentes de su microbiota intestinal en comparación con individuos sanos. Se ha observado una

reducción de la biodiversidad de *Bacteroides* y *Firmicutes phyla*, concretamente en tres subtipos como son *Faecalibacterium prausnitzii*, *Butyricoccus pullicaecorum* y *Roseburia hominis*. Por el contrario, presentan un aumento de algunos gérmenes patógenos de la familia de *Enterobacteriaceae* (*Escherichia coli*, *Klebsiella pneumoniae* and *Proteus mirabilis*) o *Fusobacterium* que serían los responsables de inducir esa inflamación continua.

A pesar de que en los individuos sanos la microbiota se modifica continuamente; se ha observado como existen diferencias entre la microbiota de estos individuos en comparación con los pacientes afectos de EC e incluso con los familiares de los pacientes afectos con EC; lo cual sugeriría una asociación de los factores genéticos con la microbiota. En definitiva, en los pacientes con EC hay una inestabilidad en esta microbiota incluso en periodos de remisión.

3. Factores ambientales: la incidencia de esta enfermedad es mayor en países desarrollados, por tanto, aunque la genética juega un papel importante, su combinación con éstos debe ser clave en la etiopatogenia<sup>18</sup>.
  - a. Tabaco: probablemente el tabaco sea el factor ambiental más estudiado en relación con la EII. Ya hace más de treinta años desde que se postuló la relación entre el tabaco y la EII<sup>19</sup>. Aunque todavía se desconoce el mecanismo exacto por el que el tabaco influye en la EII, se postula que las más de 3500 sustancias químicas que forman parte del tabaco, tienen efecto en la musculatura de los vasos sanguíneos. No se sabe si esto está relacionado con la patogenia de la enfermedad o bien es consecuencia de ésta. En el caso de la EC, el mecanismo no es totalmente conocido, aunque podría estar relacionado con una reducción de los niveles de óxido nítrico que provoca un proceso de arteritis, es decir, un fenómeno de alteración en la capacidad de vasodilatación asociado a fenómenos de trombosis que provocan finalmente oclusión de los vasos y microinfartos, perpetuando la isquemia y fibrosis<sup>20-22</sup>. Esta es la teoría que se postula gracias a estudios

iniciales<sup>23</sup> y que se ve reforzada por los estudios epidemiológicos más recientes. En ellos, se relaciona el hábito tabáquico con un incremento del riesgo de desarrollar EC (Hazard Ratio 1,9 con un IC 95% 1,42-2,53)<sup>24</sup>, con una enfermedad más agresiva, (concretamente de fenotipo estenosante-penetrante), con mayor número de brotes, e incluso con mayor utilización de glucocorticoides, tratamiento inmunosupresor y biológico<sup>25-27</sup>. Además los fumadores precisan más precozmente cirugías resectivas<sup>28,29</sup>. También se ha descrito un cambio en la microbiota con el cese del hábito tabáquico, con un aumento en la biodiversidad de microorganismos comensales como *Firmicutes* y disminución de patógenos como *Bacteroidetes*<sup>30</sup>.

- b. Dieta: los hábitos alimentarios se modifican continuamente por el estilo de vida y por la influencia de otros países, teniendo una importante repercusión en la microbiota intestinal.

Se han postulado muchas teorías sobre el consumo de grasas, carne roja, alimentos procesados, fibra o carbohidratos, no obstante, no se disponen de datos contrastados, lo que impide realizar unas recomendaciones sobre una dieta específica para los pacientes con EC, sobre todo centrada en la prevención de los brotes. Aun así, según lo publicado recientemente podrían ser factores protectores: el consumo de fibra procedente de frutas y verduras<sup>31</sup>, el consumo de leche o derivados lácteos<sup>32</sup> o bien, los suplementos o niveles adecuados de vitamina D<sup>33,34</sup>. Como factores de riesgo podrían ser, el consumo de carne roja y procesada<sup>35</sup>.

Hay más controversia sobre la dieta rica en grasas. En un estudio prospectivo de larga evolución no se ha observado una asociación significativa entre el consumo de grasas con el desarrollo de EC<sup>36</sup>. No obstante, los datos sobre la suplementación o la dieta rica en omega-3 es contradictoria. Se han publicado algunos casos en los que la dieta rica en pescado (fuente omega-3) tiene un efecto protector en el desarrollo de EC, o que el ratio omega-3:omega-6 elevado mantiene la remisión en pacientes con EC<sup>37</sup>. Sin

embargo, una revisión Cochrane sobre el tema concluye que la suplementación de omega 3 no es un factor protector en el mantenimiento de la remisión en EC<sup>38</sup>.

Por último, la nutrición enteral induce la remisión en la población pediátrica y parece que puede inducir cambios en la microbiota, pero todavía no existe suficiente información al respecto<sup>39</sup>.

- c. Fármacos: en un estudio reciente, parece que tanto los antiinflamatorios no esteroideos como los anticonceptivos orales están relacionado con un incremento de la susceptibilidad de padecer la enfermedad<sup>18,40</sup>. Los antibióticos, sobre todo las penicilinas utilizadas en el primer año de vida se ha observado como incrementan el riesgo 5,3 veces de padecer EC. Ello va disminuyendo conforme se utilizan en edades más avanzadas<sup>41</sup>.
- d. Otros factores: la apendicectomía y la lactancia materna se postularían como factores protectores y el estrés psicosocial como factor de riesgo para desencadenar un brote.

Mención especial requieren las infecciones entéricas que producen disrupción en la barrera entérica provocando el desarrollo de la cascada inflamatoria y la autoperpetuación en los casos de individuos genéticamente predisuestos. A pesar de relacionar a múltiples organismos, como es el caso de la *Yersinia enterocolitica*, *Listeria monocytogenesis*, *Escherichia Coli* o *Mycobacterium avium*, con la EC, no existen datos firmes que establezcan una relación causal.

#### **2.1.4. Manifestaciones clínicas.**

En general la EII puede debutar de muchas formas debido a que su expresión clínica depende fundamentalmente de la localización y del patrón de la enfermedad.

Se debe sospechar en los casos de diarrea crónica, habitualmente con una duración superior a 6 semanas, para diferenciarlo de la etiología infecciosa ya que la diarrea es la manifestación más frecuente<sup>42</sup>.

También puede cursar como dolor abdominal. En la mayoría de casos (80%) en forma de dolor abdominal crónico pero en una pequeña proporción pueden aparecer como cuadros similares a la apendicitis aguda debido a la afectación del íleon terminal<sup>43</sup>.

Hay que destacar que en ocasiones los pacientes debutan con sintomatología inespecífica, muy similar a la que presentan los pacientes con síndrome de intestino irritable. De hecho, se ha observado que los pacientes con EII presentan mayor prevalencia de sintomatología compatible con este síndrome<sup>44</sup>. Sin embargo, siempre se debe continuar el estudio cuando la sintomatología intestinal se asocia a anemia crónica<sup>45</sup>, alteración del crecimiento en el caso de niños o bien a productos patológicos en heces como sangre o moco<sup>46</sup>.

La afectación perianal se presenta en un 4-10% de los pacientes en el momento del diagnóstico ocasionando síntomas muy molestos que deterioran de forma sustancial la calidad de vida de los pacientes<sup>47</sup>.

También pueden aparecer manifestaciones sistémicas como la fiebre, que puede deberse al propio proceso inflamatorio o bien secundario a complicaciones como abscesos, perforaciones o sobreinfecciones. Es frecuente la pérdida de peso, en un 60% de los casos, debido al catabolismo acelerado en las fases activas de la enfermedad o bien a las resecciones y a los problemas en la absorción de nutrientes.

Otras manifestaciones sistémicas son las manifestaciones extraintestinales, que aparecen en 6-47% de los pacientes y en ocasiones preceden a los síntomas gastrointestinales<sup>48</sup>. Estas manifestaciones merman de forma importante la calidad de vida de los enfermos y algunas de ellas como la trombosis, pueden poner en peligro su vida. Esta sintomatología es más frecuente en los pacientes con EC que presentan afectación del colon.

Dentro de las manifestaciones extraintestinales, la afectación articular (artropatía periférica o axial, sacroileítis), es la más frecuente, alcanzando cifras de hasta el 50% en algunas series<sup>49</sup>.

La afectación cutánea<sup>50</sup> es la segunda en frecuencia, apareciendo en hasta un 14% de los casos, siendo el eritema nodoso la principal manifestación (4-6%), sobre todo en mujeres jóvenes y con afectación de colon. Otras manifestaciones que aparecen más esporádicamente son el pioderma gangrenoso o el Síndrome de Sweet. Dentro de la afectación cutánea se puede incluir la afectación oral que es muy frecuente, sobre todo, en forma de estomatitis aftosa o periodontitis.

La afectación ocular es la tercera en frecuencia, afectando a menos del 10% de los pacientes con manifestaciones extraintestinales y sobre todo a pacientes con EC. Suele aparecer como epiescleritis o uveítis<sup>51</sup>.

También puede aparecer la afectación hepática como la hepatitis granulomatosa, la amiloidosis, el hígado graso, la colangitis esclerosante primaria, la trombosis portal o el colangiocarcinoma<sup>52,53</sup>.

Además de las propias manifestaciones extraintestinales podemos destacar las complicaciones secundarias tanto al proceso inflamatorio como a los tratamientos administrados. Entre ellas, se encuentra la alteración del metabolismo óseo, osteoporosis, que afecta a casi la mitad de los pacientes con EC y que se presenta debido tanto a los tratamientos administrados como a la malabsorción secundaria a la propia inflamación crónica y a las resecciones intestinales<sup>54</sup>. Como se ha destacado también la enfermedad se puede complicar con la aparición de enfermedad tromboembólica venosa<sup>55</sup>.

Muchas manifestaciones aparecen en paralelo con el brote de actividad intestinal, mientras que otras, como la espondilitis anquilosante y la uveitis son independientes.

### **2.1.5. Criterios diagnósticos.**

En esta enfermedad resulta fundamental una anamnesis y exploración minuciosa para descartar las causas infecciosas, neoplásicas, actínicas o isquémicas que puedan justificar la sintomatología del paciente. Además, también es necesario tener en cuenta factores como

gastroenteritis previas, historia previa de intolerancias, viajes recientes, toma de antiinflamatorios o hábito tabáquico.

En estos momentos no existe un patrón oro que permita establecer el diagnóstico inequívoco de la EC ya que los hallazgos no son patognomónicos. Según el consenso de la *European Crohn's and Colitis Organisation* (ECCO), el diagnóstico de EC se basa en la combinación de cuatro grupos de criterios: clínicos, radiológicos, endoscópicos e histológicos, apoyados por datos analíticos<sup>56</sup>. La forma de presentación de la enfermedad puede ser muy heterogénea, desde sintomatología claramente intestinal a síntomas extraintestinales o incluso aparecer exclusivamente como manifestaciones sistémicas.

Ante la sospecha de EC, la primera exploración a realizar es una ileocolonoscopia con toma de biopsias de todos los tramos intestinales independientemente de la afectación endoscópica. Con esta exploración se puede evaluar tanto el colon como los últimos tramos del intestino delgado (íleon terminal). Los hallazgos endoscópicos que se pueden encontrar son:

- Afectación colon: úlceras aftoides o serpinginosas (superficiales o profundas) que proporcionan un aspecto en empedrado, estenosis, pseudopólipos e incluso algún orificio fistuloso. La afectación en este caso es parcheada, a diferencia de la CU, y con el recto generalmente indemne.
- Íleon: úlceras aftoides, ulceraciones serpinginosas o incluso estenosis.

No obstante, puede estar afectado cualquier tramo intestinal, por lo que si la sospecha es firme y la ileocolonoscopia es negativa se recomienda realizar un estudio con endoscopia digestiva alta. En los casos de sospecha de EC con estudio endoscópico convencional negativo y radiología normal, la cápsula endoscópica permite valorar directamente cambios intestinales precoces en yeyuno e íleon, no alcanzables por otras técnicas.

Otra herramienta útil en el diagnóstico o la valoración de la extensión son las técnicas radiológicas. Actualmente se utilizan fundamentalmente dos de ellas, la enterografía por resonancia magnética (entero-RM) o por tomografía axial computarizada (entero-TC), que



permiten evaluar la afectación transmural y la presencia de complicaciones asociadas como abscesos o estenosis.

En cuanto a la histología, los hallazgos tampoco son patognomónicos, pero algunos de ellos son muy sugestivos como la presencia de granulomas no caseificantes. Otros hallazgos que podemos encontrar son el infiltrado linfoplasmocitario en la lámina propia, infiltrado inflamatorio transmural, agregación linfoide o hiperplasia de criptas.

Debido a esta dificultad en el diagnóstico de la EC se desarrollaron los criterios de Lennard-Jones<sup>57</sup> (Tabla 1). Se requieren como mínimo dos criterios, siendo el histológico el más importante. Se considera EC establecida cuando se combina un criterio mayor, que es la presencia de granulomas no caseificantes en la histología, asociado a otro criterio o bien tres criterios menores (afectación transmural: fístula o absceso, afectación macroscópica discontinua, fibrosis/estenosis, agregados linfoides, inflamación discontinua en la histología)<sup>58</sup>.

A pesar de la combinación de criterios, hasta en un 5% de los casos se modifica el diagnóstico a lo largo de la evolución, pasando de EC a CU o viceversa. De momento el estudio genético no está recomendado de rutina para el diagnóstico o el manejo de la EC.

Tabla 1. Criterios de Lennard-Jones (Adaptada de *Lennard-Jones et al.*<sup>57</sup>).

	<b>Clínica o Endoscopia</b>	<b>Radiología</b>	<b>Biopsia</b>	<b>Cirugía</b>
<i>Lesión digestiva alta</i>	+	+	+	+
<i>Perianal</i>	+		+	+
<i>Distribución segmentaria</i>	+	+	+	+
<i>Lesión transmural</i>				
- <i>Fisura</i>		+		+
- <i>Absceso</i>	+	+		+
- <i>Fístula</i>	+	+		+
- <i>Estenosis</i>	+	+		+
<i>Histología</i>				
- <i>Úlcera</i>			+	+
- <i>Agregado linfoide</i>			+	+
- <i>granulomas</i>			+	+

A pesar de ser criterios utilizados desde hace años y ser los recomendados por la ECCO, no han sido validados y en muchas ocasiones son demasiado estrictos para realizar un diagnóstico precoz de la enfermedad. Debido a ello, recientemente se ha desarrollado un índice basado en estos criterios para realizar un diagnóstico precoz de la enfermedad, aunque todavía debe ser validado<sup>59</sup>.

### 2.1.6. Clasificación.

La complejidad en la etiopatogenia de la EC va a dar lugar a distintos patrones de comportamiento con su consiguiente variabilidad en cuanto a la expresión clínica, por lo que la clasificación fenotípica resulta fundamental.

En 1975 es cuando se estableció la primera clasificación anatómica de la EC <sup>60</sup>. Posteriormente, se observó cómo el patrón de comportamiento resultaba fundamental a nivel pronóstico, ya que el patrón penetrante en comparación con el inflamatorio presentaba un curso más agresivo con mayor número de intervenciones quirúrgicas<sup>61</sup>.

Tras estos estudios en los que se mostraron la importancia de la localización y el patrón de comportamiento, en el año 2000 se elaboró la primera clasificación de la EC, la Clasificación de Viena (Tabla 2), que incluyó la localización de la enfermedad, el patrón de comportamiento y la edad al diagnóstico como variable clínica<sup>62</sup>.

Tabla 2. Clasificación de Viena EC ( Adaptada de *Gasche et al* <sup>62</sup>).

<b>Edad al diagnóstico</b>	<b>Localización</b>	<b>Comportamiento</b>
<b>A1&lt;40 años</b>	L1: ileal	B1: ni estenosante ni penetrante
<b>A2≥40 años</b>	L2: cólica	B2: estenosante
	L3: ileocólica	B3: penetrante
	L4: Gastrointestinal alta	

Posteriormente en 2005 se hizo una modificación de la misma, apareciendo la clasificación de Montreal (Tabla 3)<sup>63</sup>. En esta nueva clasificación se subdividió la edad al diagnóstico y la localización gastrointestinal alta se añadió a cualquiera de las localizaciones intestinales, al igual que la afectación perianal se añadió a cualquier patrón de comportamiento. Actualmente es la clasificación recomendada en las guías de la ECCO.

Tabla 3. Clasificación de Montreal (*Adaptada de Silverberg et al*<sup>63</sup>).

Edad al diagnóstico (A)	Localización (L)	Comportamiento (B)
A1 ≤16 años	L1: ileal	B1: inflamatorio
A2: 17-40 años	L2: cólica	B2: estenosante
A3 >40 años	L3: ileocólica	B3: penetrante
	Añadir Gastrointestinal alta	L4: Añadir p (afectación perianal) a cualquiera de las anteriores

- De L1 a L3, se añade L4 cuando coexiste afectación gastrointestinal alta.
- En el patrón de comportamiento se añade "p" en caso de afectación perianal.

Prácticamente el 50% de los pacientes presentan una localización ileocecal, un 30% localización colónica y otro 20% afectación exclusiva ileal<sup>7,64-67</sup>. En la mayoría de los casos la localización no se modifica con el transcurso de la enfermedad, no como ocurre con el patrón de comportamiento. Al diagnóstico, los pacientes suelen debutar con un patrón inflamatorio, sin embargo, conforme progresa la enfermedad, se postula que hasta un 20-30% de los pacientes a los 5 años pueden desarrollar complicaciones como estenosis o fístulas, pasando a un patrón estenosante o fistulizante<sup>64,68</sup>, lo que le confieren una mayor agresividad debido a la necesidad de intervenciones quirúrgicas o tratamientos más agresivos.

### 2.1.7. Valoración de la actividad inflamatoria

La EC es una enfermedad crónica que precisa un seguimiento muy estrecho. En los últimos años el objetivo prioritario es alcanzar la curación mucosa, debido a que esto predice un mejor pronóstico a largo plazo. Por ello, resulta fundamental identificar de la forma más precoz posible la actividad inflamatoria.

Para evaluar esta actividad se pueden utilizar criterios clínicos, endoscópicos, radiológicos o histológicos. Según estos criterios, a lo largo de los años, se han desarrollado varios índices para cuantificar esa actividad inflamatoria, y aunque se suelen utilizar en el contexto de ensayos clínicos para definir lo que se considera respuesta, remisión, recaída o éxito terapéutico; es importante conocer tanto sus ventajas como sus limitaciones.

Disponemos, por tanto, de índices clínicos, endoscópicos, radiológicos e histológicos, además de múltiples marcadores no invasivos tanto serológicos como fecales para cuantificar la actividad inflamatoria.

### *Índices clínicos.*

Durante muchos años la valoración de la actividad se ha equiparado a la gravedad clínica; por ello se elaboraron índices clínicos para cuantificar la actividad inflamatoria y poder discernir entre enfermedad inactiva o activa, y dentro de ésta, intentar gradar la actividad en leve, moderada o grave.

Para valorar la actividad en el patrón inflamatorio de la EC disponemos de tres índices, los índices de Best, van Hess o Harvey-Bradshaw.

El índice de Best o CDAI (*Crohn's Disease Activity Index*)<sup>69</sup>, es el más utilizado sobre todo en el contexto de ensayos clínicos, además de ser el único validado. Incluye siete variables clínicas (número de deposiciones, dolor abdominal, estado general, presencia de manifestaciones extraintestinales, antidiarreicos orales, presencia de masa abdominal y peso) y sólo un parámetro analítico (hematocrito). Por tanto, presenta una elevada subjetividad, y refleja fundamentalmente la gravedad clínica. Además de la subjetividad y la variabilidad en el cálculo, otro de los inconvenientes es el tiempo de cálculo, ya que precisa de 7 días de recogida de datos, por lo que no se utiliza en la práctica clínica diaria.

La puntuación oscila desde 0 hasta 600. Si la puntuación es menor a 150 se considera enfermedad inactiva, 150-220 brote leve, de 220-450 brote moderado y más de 450 brote grave.

El índice de Harvey-Bradshaw<sup>70</sup> incluye cinco variables clínicas; el estado general, dolor abdominal, número de deposiciones líquidas, presencia de masa abdominal y complicaciones. La actividad se clasifica según sea la puntuación: menos de 5 puntos inactivo, de 5-7 brote leve, de 8-16 brote moderado y más de 16 puntos brote grave. Al ser diseñado en un estudio prospectivo y correlacionarse con el CDAI, es el que se utiliza en la práctica clínica diaria por su simplicidad<sup>71</sup>.

El índice de Van Hess<sup>72</sup> se desarrolló unos años más tarde. Incluye 8 variables clínico-analíticas (albúmina, velocidad de sedimentación globular, sexo, temperatura, consistencia de las heces, masa abdominal, manifestaciones extraintestinales e índice de masa corporal) y, además, la presencia de resecciones intestinales. No obstante, su cálculo es complejo y además su diseño fue realizado para un estudio retrospectivo y con una muestra pequeña de pacientes.

La enfermedad perianal se valora por separado y para ello se desarrolló otro índice de actividad, el PDAI (Perianal Disease Activity Index)<sup>73</sup> y el índice de Present.

Sin embargo, posteriormente se observó como en la EC la actividad inflamatoria y la gravedad clínica no siempre son concordantes. En muchas ocasiones la sintomatología de los pacientes, incluida en las variables utilizadas en los índices previamente mencionados para valorar la actividad inflamatoria, no refleja la inflamación propiamente dicha sino complicaciones secundarias a la misma como las estenosis o fibrosis, o bien secundarias a la cirugía, como sobrecrecimiento bacteriano o síndrome del intestino corto.

Por ello en los últimos años, este tipo de índices han sido sustituidos o complementados por los índices endoscópicos y/o marcadores serológicos o fecales, con la intención de obtener datos más objetivos de inflamación.

### *Índices endoscópicos.*

Realmente la EC puede afectar a cualquier segmento intestinal y la endoscopia tiene un papel fundamental no sólo en el diagnóstico de la enfermedad sino también en valorar la

actividad inflamatoria durante el seguimiento de ésta. Debido a su importancia, el papel de la endoscopia digestiva, incluyendo los índices endoscópicos, lo desarrollaremos en otro apartado de forma mas detallada.

### *Índices radiológicos.*

El avance continuo y constante en las técnicas de imagen las han convertido en otra de las herramientas de las que se disponen en la actualidad para valorar el tracto digestivo. Realmente las exploraciones radiológicas son complementarias a las endoscópicas y tienen un interés creciente en la EC debido a su afectación transmural, a la capacidad de discernir actividad inflamatoria de fibrosis, a la posibilidad de valorar zonas no accesibles por endoscopia y también a la posibilidad de diagnosticar complicaciones como las fístulas, abscesos o estenosis.

Actualmente existen tres exploraciones radiológicas para valorar la actividad inflamatoria: la ultrasonografía, la entero-TAC y la entero-RM. La ultrasonografía se ha desarrollado mucho en los últimos años, sobre todo gracias a la incorporación del contraste y de la elastografía, convirtiéndose en una alternativa prometedora que además es inocua. Sin embargo, sigue siendo una técnica explorador-dependiente, que no permite valorar la totalidad del intestino, siendo, por tanto, imposible valorar todas las complicaciones extraintestinales y de la que no disponemos de índices objetivos para evaluar la actividad inflamatoria y la gravedad de las lesiones. Por ello, aunque todas las técnicas radiológicas son útiles, se recomienda el uso de la entero-TC o entero-RM, ya que detectan inflamación en más del 90% de los casos<sup>74</sup>.

La entero-TAC, tiene como mayor desventaja la exposición a la radiación, pero es una prueba bien tolerada y muy accesible en la mayoría de los centros. La actividad inflamatoria la valora utilizando diferentes parámetros: el engrosamiento de la pared intestinal, el realce mucoso, la estratificación de la pared, la vascularización (el signo del peine provocado por el engrosamiento de los vasos rectos a la luz intestinal)<sup>75</sup> y el tamaño y realce de los ganglios linfáticos<sup>76</sup>. Chiorean et al publicaron un estudio en el que

compararon esos parámetros radiológicos con la pieza quirúrgica. Los resultados mostraron que los hallazgos radiológicos se correlacionaban con los histológicos, de hecho, la entero-TAC permitió diferenciar de forma adecuada las lesiones inflamatorias de las fibroestenóticas con una exactitud diagnóstica del 77% y 79% respectivamente<sup>77</sup>.

La entero-RM se ha posicionado durante los últimos años como la alternativa más eficaz frente a la entero-TAC tanto para valorar la extensión como la actividad inflamatoria como la presencia de complicaciones, debido fundamentalmente a la ausencia de radiación ionizante. La desventaja es que es una técnica costosa, que precisa de un entrenamiento para su realización, así como unos protocolos bien diseñados y que no está disponible en todos los centros.

Con la entero-RM disponemos de varios índices para evaluar la actividad y gravedad de las lesiones. De ellos, el más utilizado es el índice MaRIA (*Magnetic Resonance Index of Activity*) desarrollado en el 2009 por un grupo español<sup>78</sup> validado en dos cohortes de pacientes con buena correlación con el CDEIS (*Crohn Disease Endoscopic Index Activity*)<sup>79</sup>, y que, además permite valorar la respuesta a la terapia administrada, siendo incluso capaz de detectar la curación mucosa<sup>80</sup>.

Recientemente se ha publicado un estudio prospectivo que compara el índice MaRIA con los índices endoscópicos, el CDEIS y SES-CD (*Simple Endoscopic score Crohn Disease*), con el objetivo de seleccionar a los pacientes para su inclusión en ensayos clínicos terapéuticos<sup>81</sup>. En este estudio en el que se incluyen sólo 20 pacientes, se observa como existe una correlación significativa entre el índice radiológico y los índices endoscópicos (CDEIS 0,64 y SES-CD 0,71), sin embargo, no se obtiene una buena correlación con los índices clínicos, CDAI (0,30). Aun así, se observan diferencias según el tramo intestinal afecto, con peor concordancia en tramos proximales.

Disponemos de otros índices radiológicos como el *Crohn's disease MRI Index Score* (London)<sup>82</sup>, desarrollado y validado en 2012. El *Magnetic Resonance Enterography Global Score* (MEGS) desarrollado a partir del anterior, validado y capaz de valorar la respuesta a los fármacos biológicos, con buena concordancia interobservador<sup>83</sup>, y otros dos como el Nancy Score<sup>84</sup> o el Índice de Lémann<sup>85,86</sup> menos utilizados en la actualidad.



El índice de Sailer es otro índice radiológico propuesto para valorar la recurrencia en los casos de EC intervenida. Lo desarrollaremos en el siguiente apartado.

### *Marcadores no invasivos bioquímicos:*

Las técnicas endoscópicas y radiológicas son muy útiles para valorar la actividad inflamatoria, sin embargo, en la práctica clínica, son exploraciones invasivas, costosas e inviables de realizar de forma continua para monitorizar la actividad inflamatoria. Por ello, se ha puesto especial atención en el desarrollo de tests no invasivos que permitan realizar una evaluación objetiva y una monitorización del paciente.

Se han estudiado anticuerpos como los ASCAS (*anti-Saccharomyces Cerevisiae Antibodies*); no obstante, no está generalizado su uso de forma rutinaria debido a que parecen más útiles en distinguir EC de CU que para valorar la actividad inflamatoria *per se*. También disponemos de marcadores serológicos como son la proteína C reactiva (PCR) o la velocidad de sedimentación globular. La PCR se sintetiza en el hígado y se eleva rápidamente ante un proceso inflamatorio disminuyendo precozmente cuando se resuelve (vida media de 19 horas). Aunque una PCR elevada se asocia con una mayor respuesta al tratamiento biológico, se ha observado que hay pacientes con brote que tienen una PCR normal. La velocidad de sedimentación globular en cambio, tiene una vida media mayor y además se ve afectada por la edad, el hábito tabáquico y otros factores como la anemia, por lo que su uso en la práctica diaria es limitado. Ambos marcadores no son específicos de inflamación intestinal ni tampoco se correlacionan con la actividad endoscópica<sup>87</sup>.

Por ello, con el tiempo se desarrollaron los marcadores fecales, que, aunque son específicos del tracto digestivo, no lo son de la EC, ya que pueden estar elevados en infecciones, neoplasias o toma de fármacos como los antiinflamatorios no esteroideos. Dentro de ellos, la calprotectina fecal (CF), es la más estudiada y la más utilizada. Se trata de una proteína que aparece en heces de forma proporcional a la migración de neutrófilos hacia el tracto intestinal, constituyendo un marcador precoz y directo de inflamación<sup>88</sup>. Además, es muy estable a temperatura ambiente, lo que facilita su recogida y su procesamiento.

Se ha demostrado su utilidad para valorar el brote, la respuesta al tratamiento y la recidiva<sup>89-93</sup>. También se ha observado una buena correlación con los índices endoscópicos<sup>94,95</sup>. El único inconveniente, es que se desconoce el punto de corte óptimo ya que varía en los estudios y según el contexto<sup>89</sup>.

## 2.2 Cirugía en la Enfermedad de Crohn: Recurrencia.

### 2.2.1. Introducción.

En el mismo momento del diagnóstico y a lo largo de la evolución de la enfermedad, hasta el 80% de los pacientes precisan en algún momento una intervención quirúrgica, debido a la refractariedad a los tratamientos convencionales o a la aparición de complicaciones como estenosis intestinales, abscesos, plastrones o fístulas.

En estudios iniciales, cuando no se disponía de inmunomoduladores ni fármacos biológicos, se observó cómo la probabilidad de precisar una resección intestinal se incrementaba con los años de evolución<sup>96,97</sup>. En los primeros estudios poblacionales publicados la probabilidad de cirugía se situaba entre el 14-24% al año, 27-35% a los 5 años y 38-50% a los 10 años<sup>97,98</sup>. Pero no sólo eso, sino que los pacientes intervenidos también presentaban una probabilidad acumulada de cualquier tipo de resección de hasta el 16%, 25% y 35% a los 5,10 y 15 años de la primera intervención<sup>98</sup>.

Probablemente debido a los avances en el conocimiento de la etiopatogenia y en las posibilidades terapéuticas en los últimos años parece que la tendencia está cambiando. En uno de los estudios poblacionales más recientes la probabilidad acumulada de precisar cirugía, tanto en el momento del diagnóstico como en la evolución se ha reducido, al 19% y 25% respectivamente<sup>68</sup>. Y no sólo eso, sino que en un metaanálisis reciente se observa la tendencia decreciente de las necesidades quirúrgicas estratificado en función del año del diagnóstico y la aparición de fármacos inmunomoduladores (Figura 1)<sup>99</sup>.

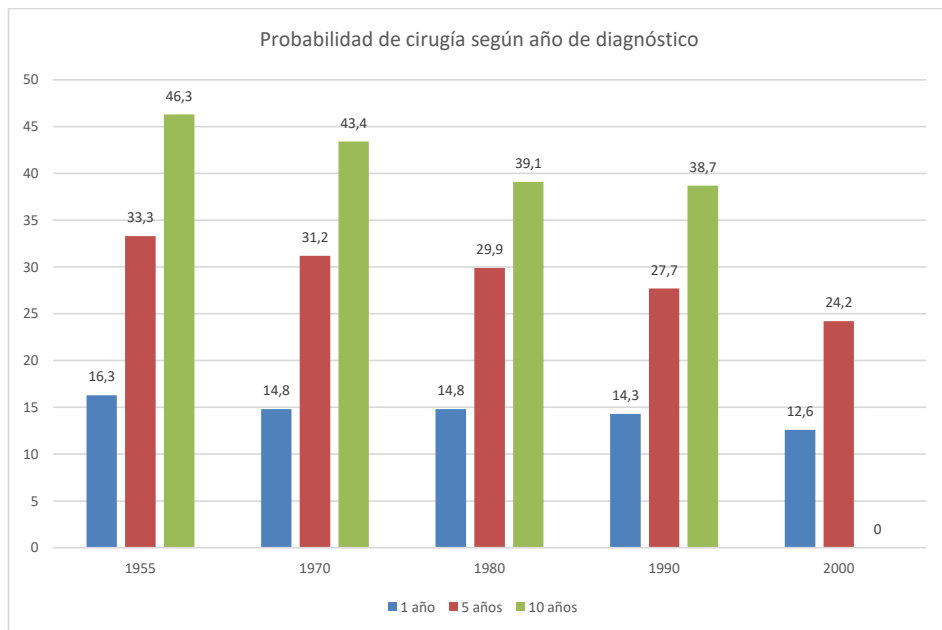


Figura 1. Probabilidad de intervención quirúrgica en EC dependiendo del año de diagnóstico y tratamientos disponibles. *Adaptada de Frolkis et al.*<sup>99</sup>.

Aunque con el desarrollo de nuevas dianas terapéuticas las probabilidades de precisar una intervención quirúrgica son menores, la cirugía sigue siendo una opción terapéutica en la EC y en sus complicaciones.

### 2.2.2. Recurrencia: definiciones.

Al tratarse de una enfermedad crónica y difusa de todo el tracto gastrointestinal, la cirugía no es curativa, reapareciendo la enfermedad en zonas previamente no afectas macroscópicamente como suele ser la propia anastomosis o regiones proximales a ésta. Este concepto es lo que se acuña con el término recurrencia postquirúrgica (RPQ). Por

tanto, podemos establecer varias definiciones según nos refiramos a recurrencia histológica, morfológica o clínica. La RPQ histológica se refiere a la aparición de cambios inflamatorios a nivel microscópico sin lesiones visibles; la RPQ morfológica se refiere a la aparición de lesiones intestinales detectadas por endoscópica (RPQ endoscópica) o por técnicas de imagen (RPQ radiológica). La RPQ clínica es la sintomatología atribuible a la presencia de lesiones morfológicas, y, por último, la RPQ quirúrgica, que sería la necesidad de nueva cirugía por la presencia de lesiones.

En general, se prefiere utilizar el término de RPQ endoscópica, ya que se ha observado como la aparición de lesiones endoscópicas precede a la identificación de lesiones morfológicas en técnicas de imagen, así como a la aparición de síntomas.

El primer estudio en el que se describió la RPQ fue el publicado por Rutgeerts en 1984<sup>100</sup>. En él, se realizaba una endoscopia al año de la intervención (resección ileal o ileocólica con anastomosis) en 114 pacientes y observaron que hasta un 72% presentaban lesiones macroscópicas en la ileoscopia, localizadas en los primeros centímetros del neóleon y en la propia anastomosis. Con posterioridad, el mismo grupo elaboró un índice, el Índice de Rutgeerts para clasificar la gravedad de las lesiones visualizadas y su implicación en el pronóstico de los pacientes. Así, determinaron que la presencia de lesiones extensas en el neóleon a los pocos meses de la cirugía predecía una peor evolución con desarrollo temprano de los síntomas, posibles complicaciones e incluso necesidad de reintervención<sup>101</sup>.

En un estudio publicado recientemente, aproximadamente al año de la intervención, el 75% de los pacientes presentan recurrencia endoscópica y un 25% sintomatología asociada<sup>102</sup>. A los 5 años las tasas ascienden discretamente hasta el 90% y 40% respectivamente.

### **2.2.3 Historia natural: Recurrencia post-quirúrgica.**

La etiopatogenia y por tanto los factores de riesgo de la EC no son del todo conocidos, y en el caso de la RPQ ocurre lo mismo debido al carácter retrospectivo de la mayoría de los estudios, siendo muy complicado establecer conclusiones y asociaciones definitivas.

El proceso biológico comienza pocos días después de la intervención quirúrgica, cuando se produce un incremento de las células inflamatorias alrededor de la zona de anastomosis<sup>103</sup>. Parece que este fenómeno inflamatorio se produce al reestablecerse el tránsito intestinal con el contacto de las bacterias con la mucosa intestinal, pero no debe ser el único factor puesto que en pacientes con ileostomía definitiva también se observa RPQ endoscópica<sup>104</sup>.

Precisamente en ese proceso biológico, la microbiota intestinal tiene un papel fundamental y está teniendo gran protagonismo en los últimos años. En el postoperatorio inmediato el neoíleon se ve colonizado por una gran cantidad de microorganismos procedentes tanto del íleon como del colon. A pesar de no disponer de demasiados datos al respecto, en estudios iniciales se observó cómo la reducción de *Firmicutes* (*C.coccoides*) y *Fecalibacterium prausnitzii* y el incremento de los *Escherichia coli* y *Enterococci* se asociaba a una recurrencia precoz de la enfermedad (en menos de 6 meses)<sup>105</sup>. Además, también se observó un incremento de los *Bacteroides* y las *Fusobacterias*<sup>106</sup>.

Recientemente estos datos han sido confirmados por otros dos estudios. Parece que la reducción de los *F prausnitzii* es prácticamente una constante en la recurrencia precoz, sin embargo, los *Bacterioides* se encuentran presentes tanto en pacientes con recurrencia como en remisión. Pero no sólo eso, hay que tener en cuenta que la microbiota no sólo se modifica cuando hay RPQ sino que cambia en función de la gravedad de la misma<sup>107-109</sup>.

También existen algunos factores de riesgo clínicos relacionados con la necesidad de cirugía, y, por tanto, con la recurrencia. Entre ellos destacan, el diagnóstico precoz de la enfermedad (<30 años), el patrón estenosante o fistulizante, la localización gastrointestinal alta, la afectación ileal extensa o ileocólica<sup>110</sup>, que haya precisado más de 2 cirugías<sup>102,111</sup> o que presente enfermedad perianal.

De los factores ambientales relacionados con la recurrencia, el tabaco es uno de los más estudiados y el que se ha corroborado en múltiples estudios. Los pacientes fumadores tienen más riesgo de intervención quirúrgica por presentar un curso de enfermedad más agresiva. En un metaanálisis que recoge 16 estudios se observó cómo los fumadores tenían mayor probabilidad de presentar una recurrencia temprana de la enfermedad con una OR 2.07 en comparación con los no fumadores<sup>112</sup>. Sin embargo, este efecto es reversible. El

cese del tabaquismo está relacionado con una reducción del riesgo con una Hazard Ratio 1,35 con un IC 95% (1,05-1,73), así como con la menor utilización de corticoesteroides y terapias inmunosupresoras<sup>113,114</sup>.

En cuanto a la cirugía, la resección ileocecal es la intervención más frecuente. Ni el tipo de cirugía, el tipo de anastomosis o la longitud de los márgenes de resección están relacionados directamente con la aparición de recurrencia<sup>115</sup>. No obstante, la presencia de plexitis mientérica sí que parece ser uno de los factores directamente relacionados con la presencia de una recurrencia precoz de la enfermedad, prácticamente al mes de la intervención, y que además tiene implicaciones directas con la instauración de algunos tratamientos<sup>116,117</sup>.

#### **2.2.4. Valoración de la recurrencia: métodos disponibles.**

Tras la cirugía, uno de los objetivos del seguimiento de los pacientes es la detección precoz de la RPQ endoscópica instauración del tratamiento más adecuado. Actualmente lo recomendado según la guía de la ECCO es la realización de una ileocolonoscopia a los 6-12 meses de la intervención<sup>118</sup>, sin embargo, se desconoce si es necesario realizar endoscopias de seguimiento o cada cuánto tiempo realizarlas, teniendo en cuenta que son exploraciones invasivas e incómodas para el paciente. Por ello, al margen de las exploraciones endoscópicas, se han evaluado otras herramientas que permitan una monitorización continua de los pacientes de forma no invasiva.

Para valorar la RPQ disponemos de varias herramientas. En primer lugar, se encuentran las exploraciones endoscópicas. Entre ellas la ileocolonoscopia es la técnica utilizada como patrón oro para determinar la presencia y la gravedad de las lesiones en neóleon. Como se ha comentado, en 1990, Rutgeerts et al<sup>101</sup>, desarrollaron un índice para valorar la actividad inflamatoria explorando los últimos 15-20 cm del neóleon al año de la intervención quirúrgica. Este score puntuaba de i0, ausencia de lesiones, hasta i4, presencia de ulceraciones extensas en neóleon. Este apartado lo desarrollaremos posteriormente en la sección de Endoscopia en Enfermedad de Crohn.

Sin embargo, hasta en un 10% de los casos resulta imposible la intubación de neóleon mediante colonoscopia, bien por actividad o por el tipo de anastomosis. Por ello se disponen de otras técnicas endoscópicas para este fin como la cápsula endoscópica.

Desde el año 2000 disponemos de la cápsula endoscópica, que permite valorar no sólo el neóleon sino también otros tramos de intestino delgado más proximales. Se ha demostrado su utilidad tanto en el diagnóstico como en el manejo de la EC, disponiendo incluso de índices endoscópicos para valorar la actividad exclusiva de intestino delgado (Lewis score y *Capsule Endoscopy Crohns Disease Activity Index-CECDAI*). Sin embargo, sólo existen dos estudios que valoran la utilidad de la cápsula endoscópica en la recurrencia, pero con inclusión de pocos pacientes y con resultados algo contradictorios<sup>119,120</sup>. Por una parte, un estudio prospectivo francés que compara directamente la ileocolonoscopia con la cápsula endoscópica en 32 pacientes a los 6 meses de la intervención obteniendo una mayor sensibilidad en detectar lesiones en la colonoscopia que en la cápsula endoscópica, 90% vs 62-76%<sup>121</sup>. El segundo estudio, se trata de un estudio español desarrollado por nuestro grupo, que obtiene resultados opuestos; se incluyen 19 pacientes y sólo 25% de las colonoscopias detectan lesiones frente al 68% de las cápsulas, siendo además la técnica preferida por los pacientes, por su sencillez y tolerancia<sup>122</sup>. Debido a la información contradictoria y a la ausencia de un índice endoscópico claro que nos permita evaluar la gravedad de las lesiones, en las últimas guías de la ECCO, la cápsula se reserva para los casos de sospecha de EC con ileocolonoscopia normal o en casos ya diagnosticados de EC, cuando hay síntomas que no se puede explicar por otro motivo<sup>118</sup>.

Esta situación podría verse modificada debido a la aparición en el último año de la cápsula de Crohn, una cápsula de nueva generación con dos cámaras, mayor ángulo de visión y que permite realizar panendoscopia evitando el resto de las exploraciones endoscópicas. A pesar de la existencia de estudios prometedores en este campo, su papel en la recurrencia y en el algoritmo diagnóstico todavía no está establecido<sup>123</sup>.

En segundo lugar, disponemos de métodos no invasivos que, actualmente, se reservan para el seguimiento a largo plazo de los pacientes tras el primer control endoscópico.



En este grupo se encuentran tanto marcadores serológicos como la PCR, que se ha visto cómo no se correlaciona con la RPQ endoscópica<sup>124</sup>, y los marcadores fecales, la lactoferrina o la CF, siendo esta última la más estudiada y utilizada en la práctica diaria.

En cuanto a la CF, estudios prospectivos y multicéntricos han confirmado que efectivamente, valores superiores a 100 µg/g tiene un elevado valor predictivo negativo (superior al 90%) posicionándola como una herramienta útil en la detección de la recurrencia y en la monitorización de la enfermedad tras la instauración del tratamiento<sup>124-127</sup>. No obstante, el valor predictivo positivo sigue siendo pobre. Esto, añadido a que realmente desconocemos el punto de corte óptimo, hace que la recurrencia, por el momento, deba seguir siendo confirmada endoscópicamente.

Por último, disponemos de técnicas de imagen para valorar la RPQ morfológica. A pesar de la evolución constante de la radiología, realmente hay poca información sobre cuál es la mejor técnica para evaluar la recurrencia en casos que resulte imposible acceder a neóleon<sup>76</sup>.

La ecografía, que entre otros factores valora el grosor de la pared, tiene una sensibilidad en detectar recurrencia del 80%. Si se administra contraste oral las cifras pueden llegar al 86% con una especificidad del 96% y si se administra contraste intravenoso la sensibilidad puede alcanzar el 95%. Sin embargo, los datos entre los estudios son heterogéneos debido a que se trata de una exploración operador dependiente.

Otra técnica de la que disponemos es la tomografía axial computarizada (TAC), en esta técnica los datos son escasos ya que la mayoría de los esfuerzos se centran en la entero-RM, debido a la ausencia de radiaciones ionizantes.

Para la valoración de la recurrencia mediante entero-RM, se dispone del Índice de Sailer, que clasifica los hallazgos en cuatro categorías (Tabla 4). En el estudio en el que se desarrolló el índice, se sugirió que la entero-RM podía ser equivalente a la ileocolonoscopia para valorar la recurrencia con un índice kappa global de 0,67, que aumentaba a 0,84 según el punto de corte que estableció Sailer para indicación de tratamiento (M2)<sup>128</sup>.

Tabla 4. Índice de Sailer. Adaptado de *Sailer et al* <sup>128</sup>.

Índice de Rutgeerts	EnteroRM Índice MR	Definición
<b>i0</b>	MR 0	Sin hallazgos
<b>i1-i2</b>	MR1	Anormalidades mucosas menores: engrosamiento mural o discreto realce mural tras administrar contraste. Ausencia de estenosis
<b>i3</b>	MR2	Anormalidades mucosas mayores: otros grados de engrosamiento mural e hipercaptación. Estenosis sin dilatación preestenótica.
<b>i4</b>	MR3	Mismos hallazgos que MR2 añadiendo: Edema transmural con aumento de señal en T2 e hipercaptación de contraste en la grasa circundante. Estenosis con dilatación preestenótica y complicaciones mayores como fístulas, abscesos y conglomerado de asas intestinales

Por tanto, a pesar de disponer de un abanico cada vez más amplio de herramientas diagnósticas, en este momento la ileocolonoscopia sigue siendo una pieza clave en la valoración de la recurrencia de la enfermedad.

### 2.2.5. Prevención de la Recurrencia.

A lo largo de estos años se han evaluado diferentes grupos de fármacos en la prevención de la recurrencia como la mesalazina, los inmunomoduladores tiopurínicos, los fármacos biológicos y los antibióticos (metronidazol).

En el último consenso de la ECCO, sólo se recomienda la mesalazina a dosis altas en aquellos pacientes de bajo riesgo de recurrencia y una resección aislada de íleon o bien en aquellos en los que resulta imposible administrar otros tratamientos. Los antibióticos nitroimidazólicos como el metronidazol, se han demostrado eficaces en la prevención de la recurrencia si se administran durante los 3-12 meses inmediatamente después de la cirugía. No obstante, su principal limitación son los efectos secundarios tanto a corto plazo por intolerancia digestiva como a largo plazo por neurotoxicidad irreversible.

Los inmunomoduladores, azatioprina y 6 mercaptopurina son más eficaces que placebo en prevención de la recurrencia tanto endoscópica (i2-i4) como clínica, sin embargo, parece que no son tan eficaces en el subgrupo de pacientes con recurrencia severa, i3-i4<sup>129</sup>.

Por último se encuentran los fármacos biológicos, como los anti-TNF (*anti-tumoral necrosis factor*). El Infliximab (IFX) fue el primer anti-TNF disponible. En un ensayo clínico controlado con placebo a dosis de 5mg/kg cada 8 semanas se demostró su superioridad en prevención de la RPQ endoscópica. A los 18 meses de tratamiento el 51% del grupo placebo tenía RPQ endoscópica frente al 22% del grupo de IFX,  $p < 0.01$ <sup>130</sup>.

Posteriormente apareció el Adalimumab (ADA), con resultados controvertidos. En uno de los primeros estudios se demostró la eficacia del ADA en la prevención de la RPQ endoscópica frente a la azatioprina y la mesalazina en 51 pacientes con cifras de 6.3% vs 64.7% y 83.3% respectivamente<sup>131</sup>. En un subanálisis del estudio POCER, también se comparó la eficacia de las tiopurinas y ADA en pacientes de alto riesgo de recurrencia precoz. A pesar de que concluyeron que el grupo de ADA tenía menos RPQ endoscópica (21% vs 45%  $p = 0.02$ ); hay que tener en cuenta que consideraron RPQ endoscópica una puntuación  $\geq 2$ . Al considerar recurrencia grave, i3-i4, desapareció esa diferencia y la tasa de recurrencia en ambos grupos se igualó (8% vs 4%)<sup>132</sup>.

Recientemente se ha publicado otro estudio que ofrece resultados contradictorios. Se trata de un ensayo clínico aleatorizado multicéntrico doble ciego en el que se compara el ADA (dosis de inducción de 180 mg y 80mg y mantenimiento con 40mg/2 semanas) vs azatioprina (dosis 2-2.5mg/kg), ambos asociados a metronidazol (250 mg cada 8 horas) los

primeros 3 meses, en la prevención de la recurrencia a las 52 semanas. En este estudio no se observaron diferencias entre ambos fármacos, 29% vs 33%,  $p = 0.76$ <sup>133</sup>.

En varios metaanálisis y revisiones sistemáticas publicadas hasta la fecha en la que se evalúan todos los tratamientos disponibles, concluyen que la terapia biológica es superior al resto de tratamientos de los que se dispone actualmente<sup>134,135</sup>.

### *Prevención de la recurrencia:*

Todavía no está definida ni consensuada la estrategia a seguir en el manejo postquirúrgico de los pacientes con EC. De momento se disponen de dos estrategias de prevención de la recurrencia. La primera de ellas sería administrar un tratamiento empírico a todos los pacientes a las pocas semanas de la cirugía en contraposición a la terapia guiada por los resultados de la ileocolonoscopia realizada a las 6-12 meses postcirugía.

La primera estrategia la avala el estudio PREVENT que pretende valorar la eficacia del IFX en la prevención de la recurrencia. En este estudio se incluyeron 297 pacientes y se aleatorizaron a IFX 5mg/Kg peso cada 8 semanas vs placebo durante 200 semanas, obteniendo menores tasas de RPQ endoscópica en el grupo de IFX definida con  $i \geq 2$  en comparación con el grupo placebo (22.4% vs 51.3%  $p < 0.001$ ), sin diferencias en la RPQ clínica<sup>130</sup>.

En el estudio POCER se compararon estrategias de seguimiento en pacientes con EC intervenida. El estudio evaluó el papel de la endoscopia realizada a los 6 meses postcirugía en la monitorización del tratamiento. En este estudio se aleatorizaron a los pacientes en dos ramas, una de ellas, la rama activa en la que se realizó una colonoscopia a los 6 meses y se intensificó el tratamiento en los casos de RPQ endoscópica ( $i \geq 2$ ) y la otra rama, estándar, en la que no se realizó colonoscopia durante el seguimiento y la decisión terapéutica se basó en datos clínicos y serológicos. Además, el tratamiento de los pacientes estaba condicionado según el riesgo de RPQ endoscópica. A los pacientes de alto riesgo se les administró metronidazol combinado con un inmunomodulador, o ADA en caso de intolerancia a las tiopurinas, y a los de bajo riesgo sólo se les administró metronidazol. El

objetivo primario fue la RPQ endoscópica a los 18 meses. Los pacientes de la rama activa, a los que se les ajustó el tratamiento según los datos de la colonoscopia a los 6 meses, obtuvieron menor tasa de RPQe en comparación con la rama estándar, 49% vs 67%, ( $p = 0,03$ ), demostrando que la realización de la endoscopia permitió un ajuste de la medicación reduciendo de forma notable la RPQ endoscópica (18%)<sup>136</sup>. Además, éste es el único estudio que estratificó a los pacientes según el riesgo. Aunque los pacientes de alto riesgo obtuvieron más RPQ endoscópica que los de bajo riesgo (50% vs 70%) no alcanzaron la significación estadística ( $p = 0.04$ ).

Por otra parte, en un subanálisis del estudio anterior que incluyó a pacientes de alto riesgo de RPQ endoscópica, la recurrencia fue menor en el grupo de ADA vs AZA de forma significativa.

Otra opción en el manejo postoperatorio es esperar a la existencia de RPQ endoscópica para instaurar el tratamiento más adecuado. Con esta estrategia no se dispone de tantos datos. En un estudio se comparó la azatioprina y el infliximab en el tratamiento de la RPQ endoscópica precoz. Se incluyeron a 26 pacientes con RPQ endoscópica ( $i > 2$ ), se dividieron en tres brazos, IFX con mesalazina, azatioprina con mesalazina o bien mesalazina en monoterapia, y se observó cómo el brazo de IFX mejoraba de forma sustancial el score endoscópico comparado con el resto de tratamientos (75% vs 38% en el grupo AZA o 0% en el de mesalazina)<sup>137</sup>. Recientemente, en otro estudio prospectivo controlado y aleatorizado que incluyó a 63 pacientes, se aleatorizó a los pacientes a tratamiento con tiopurinas a las 2 semanas de la cirugía o bien inicio de tiopurinas una vez constatada la RPQ endoscópica ( $i \geq 2$ ) mediante colonoscopia a los 6-12 meses. En este estudio, no hubo diferencias en ambas ramas con un 58% de RPQ endoscópica en el grupo de pacientes sin tratamiento en comparación con 50% en el grupo tratado con azatioprina<sup>138</sup>.

Con la evidencia disponible hasta ahora y sus limitaciones parece que la estrategia más aceptada tanto a nivel nacional como internacional es la estratificación de los pacientes según el riesgo de RPQ endoscópica y en función de ello y las características basales de los mismos instaurar un tratamiento preventivo ajustándolo posteriormente según los datos endoscópicos de los que se dispongan durante el seguimiento<sup>139,140</sup>.

Recientemente se ha publicado una revisión de la *American Gastroenterological Association* sobre los aspectos más controvertidos sobre recurrencia<sup>141</sup>.

Uno de ellos se relaciona precisamente sobre cuándo iniciar el tratamiento preventivo de la RPQ endoscópica, y aunque la evidencia es baja, se recomienda iniciar la profilaxis inmediatamente tras la cirugía excepto en aquellos pacientes de bajo riesgo, en los que la profilaxis basada en los hallazgos endoscópicos sería lo más razonable, dado el riesgo-beneficio del tratamiento (Tabla 5).

Por otra parte, en cuanto al fármaco que se debe utilizar, las tiopurinas y/o anti-TNF son los fármacos de preferencia, quedando la elección de uno de ellos según las características de los pacientes. La pauta de antibióticos se reserva para aquellos pacientes de bajo riesgo y otros fármacos como la budesonida, los probióticos o bien la mesalazina no tienen papel en este escenario.

Además, también se recomienda el control endoscópico a los 6-12 meses, tanto para los pacientes que reciben profilaxis como para aquellos que no la reciben, con el fin de valorar la existencia de lesiones y ajustar el tratamiento. Aunque lo recomendado es intensificar el tratamiento en caso de existir RPQ endoscópica, en aquellos casos con lesiones menos graves, i2, podría considerarse el seguimiento endoscópico<sup>141</sup>.

Tabla 5. Riesgo de recurrencia clínica y endoscópica de la EC sin tratamiento en función de los factores de riesgo. *Adaptada de Nguyen et al*<sup>141</sup>.

	<b>Características</b>	<b>Riesgo RPQ clínica tras 18 meses cirugía</b>	<b>Riesgo RPQ endoscópica tras 18 meses cirugía</b>
<b>Bajo riesgo</b>	Mayores de 50 años. No fumadores Cirugía por estenosis corta fibrosa	20%	30%
<b>Alto riesgo</b>	Duración más 10 años Jóvenes (menores 30 años) Fumadores Más de 2 cirugías previas por enfermedad fistulizante con o sin enfermedad perianal.	50%	80%

Según esta revisión y las conclusiones que se obtienen, se ha propuesto un algoritmo como guía para elegir la opción terapéutica de los pacientes con EC ileocólica intervenidos con anastomosis ileocólica para valorar la recurrencia (Figura 2), muy similar al algoritmo proporcionado por el Grupo Español de Trabajo de EC y CU (GETECCU) (Figura 3).

#### 2.2.6. Seguimiento de la EC.

Tras la intervención quirúrgica se recomienda realizar un control endoscópico a los 6-12 meses, sin embargo, el seguimiento posterior no queda establecido, ni qué método utilizar ni tampoco con qué frecuencia. Parece que los marcadores fecales, concretamente la CF se postula como la mejor herramienta para el control evolutivo a largo plazo, realizando determinaciones cada 4-6 meses. La realización de colonoscopia o entero-RM, en caso de no ser accesible el neofleon por endoscopia, se reserva para cuando los valores de la misma sean superiores a 100 µg/gr en al menos dos determinaciones, ante sintomatología con

elevación de RFA o bien si existen datos ecográficos de recurrencia<sup>139</sup>. GETECCU también propone un algoritmo de actuación en los pacientes con EC intervenida, tanto para control inicial como para seguimiento (Figura 3).

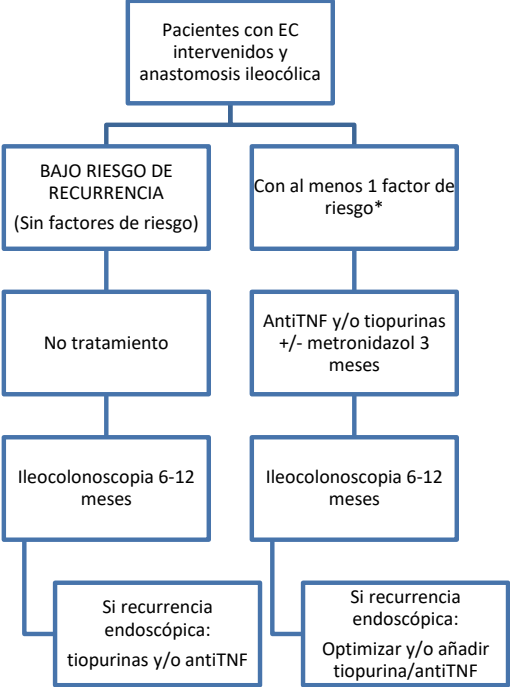
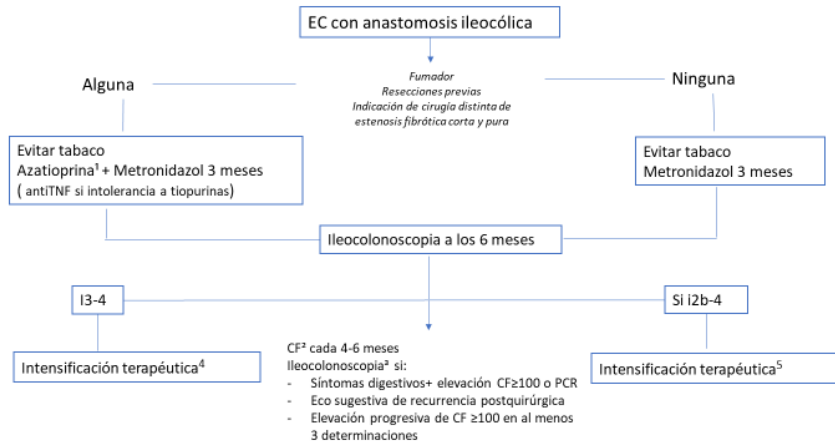


Figura 2: Algoritmo de seguimiento de pacientes con EC intervenida según la American Gastroenterological Association. *Adaptado de Regueiro et al.*<sup>140</sup>.





<sup>1</sup> antiTNF en monoterapia o en terapia combinada de forma individualizada.

<sup>2</sup> eco abdominal como alternativa a la CF.

<sup>3</sup> enteroRM en caso de ileocolonoscopia técnicamente imposible o rechazo del paciente.

<sup>4</sup> iniciar antiTNF en caso de tratamiento con tiopurinas; intensificación o cambio de antiTNF si tratamiento con el mismo.

<sup>5</sup> iniciar tiopurinas en caso de no tratamiento, mesalazina o metronidazol; iniciar antiTNF en caso de intolerancia a tiopurinas.

Figura 3. Algoritmo de seguimiento de EC intervenida según GETECCU. *Adaptado de Domenech et al*<sup>139</sup>.

### 2.3. Endoscopia en la EC.

La endoscopia tiene un papel decisivo tanto en el diagnóstico como en el seguimiento de los pacientes. La curación mucosa se ha asociado a un mejor pronóstico y por ello se ha producido un cambio sustancial en el objetivo del tratamiento de la EII. Actualmente no se busca la ausencia de síntomas sino la ausencia de lesiones, por ello, los índices de actividad endoscópicos están siendo los protagonistas no sólo de los ensayos clínicos realizados recientemente sino en la práctica clínica diaria, como un método objetivo de evaluar la actividad inflamatoria y ayudar en la toma de decisiones clínicas y monitorización del tratamiento. Además, también constituyen un lenguaje estandarizado en los informes endoscópicos para evitar la ambigüedad en las descripciones.

En la endoscopia como herramienta para evaluar la actividad inflamatoria hay que destacar que son importantes dos puntos, por una parte, los instrumentos disponibles para realizar esa evaluación, es decir, los índices endoscópicos y por otra la persona que utiliza esos índices y realiza la lectura proporcionando un diagnóstico endoscópico, el lector. Estos dos puntos van a ser revisados por separado dada su importancia.

#### 2.3.1. Descripción de los índices endoscópicos.

En cuanto a los índices endoscópicos disponibles para determinar la actividad en la EC disponemos de tres, dos de ellos para valorar la actividad luminal: el CDEIS (1989)<sup>142</sup> y el SES-CD<sup>143</sup> y el tercero para valorar la recurrencia de la enfermedad: el índice de Rutgeerts<sup>100,101</sup>. Los dos primeros validados y con buena fiabilidad.

**CDEIS:** se desarrolló en la década de los 80 por el grupo francés de estudio de enfermedades digestivas, *French Groupe d'Etude des Affections Inflammatoires du tube Digestif* (GETAID)<sup>142</sup>. En el mismo estudio, por una parte se desarrolló éste índice basándose en las variables preseleccionadas de un estudio previo<sup>144</sup>, y por otra se validó en una muestra independiente de pacientes.

Se cuantificaron cuatro variables endoscópicas: la presencia de úlceras profundas, úlceras superficiales, la extensión de la mucosa ulcerada y de la mucosa afectada y la presencia de estenosis ulceradas o no. Todo ello en las siguientes localizaciones: recto, sigma, colon izquierdo, colon transversal, colon derecho e íleon.

La puntuación se estableció de 0 a 44 y valoró los siguientes parámetros:

- Úlceras profundas (0 ausente, 12 presente).
- Úlceras superficiales (0 ausente, 6 puntos presente).
- Superficie de la mucosa ulcerada (0-10 cm): de 0 a 10 según la longitud de la mucosa ulcerada.
- Superficie de la mucosa afectada (0-10 cm): de 0 a 10 según la longitud de la mucosa afectada.

Además, añadió 3 puntos extra por la presencia de estenosis ulcerada y otros 3 puntos si había una estenosis no ulcerada (Tabla 6).

Se valora cada tramo intestinal y se divide por el número de segmentos intestinales explorados o valorados tanto completamente como parcialmente. Posteriormente se suman los puntos de las estenosis ulceradas o no.

Como se puede observar, es un índice complejo de desarrollar y laborioso que precisa de un entrenamiento para poder recoger adecuadamente los datos y obtener un resultado fiable, por lo que no resulta un índice práctico para la práctica diaria habitual.

Tabla 6. Descripción gráfica del CDEIS. Adaptado de *Mary et al.*<sup>142</sup>

	<b>Recto</b>	<b>Sigma/ Colon descendente</b>	<b>Colon Transverso</b>	<b>Colon ascendente</b>	<b>Ileon</b>	
Úlceras Profundas 0 ausente- 12 presente						Total 1
Úlceras Superficiales 0 ausente, 6 presente						Total 2
Superficie afecta por la enfermedad (cm)						Total 3
Superficie ulcerada (cm)						Total 4
Total 1+Total 2+Total 3+ Total 4						A
Número (n) de segmentos explorados parcial o totalmente (1-5)=						n
Total A/n=						B
Estenosis ulcerada (0 ausente, 3 presente)						C
Estenosis no ulcerada (0 ausente, 3 presente)						D
Total B+C+D=						

**SES-CD** : se desarrolló en 2004 a partir del anterior, pero más simplificado. Se escogieron los ítems endoscópicos más importantes y reproducibles, así como los que contribuían en mayor medida a los síntomas del paciente: tamaño de las úlceras, porcentaje de superficie ulcerada, porcentaje de superficie afecta no ulcerada y presencia de estenosis franqueable o no.

Se puntuó en cada tramo intestinal (recto, sigma-colon descendente, transverso, ascendente e íleon) cada parámetro de 0 a 3 según ausencia o presencia. La puntuación final varía de 0 a 60 (Tabla 7)<sup>143</sup>.

En ese mismo estudio se mostró que con los ítems elegidos, el SES-CD presentaba una buena correlación con el CDEIS con un coeficiente de Pearson de 0,88.

No se han definido puntos de corte ni tampoco una puntuación concreta que se considere curación mucosa. A pesar de ello, sí disponemos de una clasificación de gravedad: remisión 0-2, leve 3-6, moderada 7-15 y grave superior a 16.

Tabla 7: SES-CD. Adaptado de *Daperno et al*<sup>143</sup>.

Variable	0	1	2	3
<b>Tamaño úlcera (cm)</b>	ninguna	Úlceras aftosas (0,1-0,5)	Úlceras grandes (0,5-2cm)	Úlceras muy grandes (>2cm)
<b>Superficie ulcerada (%)</b>	ninguna	<10	10-30	>30
<b>Superficie afecta (%)</b>	ningún segmento afecto	<50	50-75	>75
<b>Estenosis</b>	ninguna	Una franqueable	Múltiples, franqueables	infranqueable

### Índice de Rutgeerts.

Es el único índice endoscópico disponible actualmente para evaluar la actividad inflamatoria endoscópica mediante ileocolonoscopia en pacientes intervenidos con resección ileocólica<sup>100,101</sup>.

Este índice clasifica las lesiones en 5 grupos: i0, ausencia de lesiones, i1 lesiones aftosas en número igual o menor a 5, i2 más de 5 lesiones aftosas entre mucosa estrictamente normal, áreas saltadas de lesiones más grandes o bien lesiones confinadas a <1cm de la anastomosis, i3 ileitis aftosa difusa sobre mucosa inflamada y finalmente i4 afectación

mucosa con amplias úlceras, nódulos y/o estenosis (Tabla 8). Además proporciona la definición de lo que se considera úlcera o lesión aftosa, que se trataría de úlceras pequeñas, redondeadas y con un halo eritematoso y discretamente sobreelevado a diferencia de las úlceras amplias, que las considera como úlceras alargadas, longitudinales, profundas y serpinginosas.

Es un índice que no sólo proporciona información sobre la gravedad de las lesiones en el momento de realizar la endoscopia, sino que es capaz de predecir la evolución clínica a largo plazo. A los 3 años, el 50% de i0 no presentan sintomatología en comparación con el 100% de los pacientes con i4, precisando hasta un 30% de los casos nueva resección quirúrgica<sup>101</sup>.

Por tanto, según los estudios iniciales realizados, este índice permitía valorar ambos parámetros, tanto la gravedad como la predicción de sintomatología. No obstante, siempre ha existido una ambigüedad en el punto de corte de este índice para considerar o no recurrencia y por tanto curación mucosa. Inicialmente se consideraba en los estudios que recurrencia era un score superior a i1. Sin embargo, posteriormente otros estudios consideraron recurrencia un score superior a i2, siendo lesiones significativas o graves las puntuaciones i3 e i4. Se ha visto que la evolución de la enfermedad en los scores i0 y i1 no difiere demasiado, pero este grupo en global sí que presenta una mejor evolución que el grupo con puntuaciones i3 y i4. Tanto en el mismo estudio en el que se desarrolló el índice como en estudios posteriores se pudo observar cómo el índice i2 incluía a pacientes heterogéneos, con ulceraciones exclusivamente en la anastomosis o bien con lesiones en el neoíleon, y esto le confería a este subgrupo un comportamiento variable. Prácticamente la mitad de ellos progresaban a lesiones más graves y a presencia de recurrencia clínica mientras que la otra mitad permanecían asintomáticos o con lesiones leves en la endoscopia.



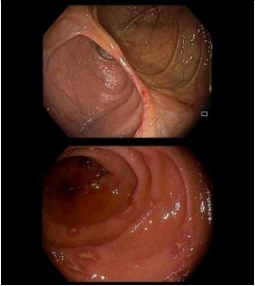


*A posteriori*, en un estudio de evaluación de azatioprina en recurrencia tanto clínica como endoscópica, se observó cómo los pacientes con úlceras en la anastomosis tenían menos probabilidad de progresar a lesiones más graves en la endoscopia y a la presencia de recurrencia clínica en comparación con los que ya presentaban lesiones en el neoíleon. Por ello, se propuso una subdivisión de este índice en i2a, que se tratarían de lesiones

confinadas en la anastomosis y cuya naturaleza podría ser isquémica, con menos probabilidad de traducirse en una recurrencia clínica, o bien i2b, en la que las lesiones se encuentran en tramos más proximales del neóleon y por tanto se relacionan con actividad de la propia enfermedad<sup>145</sup>. No disponemos de suficiente información al respecto, aunque en dos estudios retrospectivos multicéntricos no se ha observado diferencias en término de recurrencia clínica a largo plazo entre ambos grupos<sup>146,147</sup>.

Todos los estudios e incluso ensayos clínicos utilizan el índice de Rutgeerts, a pesar de no estar validado ni aprobado por la *Food and Drug Administration* (FDA). Recientemente se ha propuesto un nuevo score, POCER score, basado en la profundidad de las úlceras, número, tamaño y extensión circunferencial de las úlceras anastomóticas. Se incluyeron 118 pacientes (58 para el desarrollo del índice y 60 para su validación), se realizó colonoscopia a los 6 y 18 meses postintervención y este índice se desarrolló para predecir la RPQ endoscópica a los 18 meses. De los pacientes incluidos 19 y 30 respectivamente desarrollaron RPQ endoscópica a los 18 meses. La profundidad de las úlceras y la afectación circunferencial de las mismas en la anastomosis o en el neóleon a los 6 meses de la cirugía se relacionaron con una afectación más grave y con la necesidad de intensificar el tratamiento con una OR 1,8 (1,2-2,7 IC 95% p 0,001)<sup>148</sup>.

De momento, el índice de Rutgeerts, a pesar de no estar validado, es el aceptado y utilizado universalmente para valorar la recurrencia y tanto las guías de práctica clínica como las revisiones lo utilizan para guiar y modificar la el tratamiento a los pacientes<sup>139,141</sup>.

Tabla 8. Índice endoscópico de Rutgeerts. Adaptado de Rutgeerts et al<sup>101</sup>. Índice de Rutgeerts modificado (Adaptado de Domenech et al<sup>145</sup>)

Score	Índice de Rutgeerts	Índice Rutgeerts modificado	Imagen
<b>i-0</b>	Mucosa normal. Sin lesiones	Mucosa normal. Sin lesiones	
<b>i-1</b>	≤ 5 lesiones aftosas	≤ 5 lesiones aftosas	
<b>i-2</b>	> 5 lesiones aftosas con mucosa normal interlesional o áreas salteadas de lesiones de mayor tamaño o úlceras a <1cm de la anastomosis	i2a- lesiones en la anastomosis i2b- > 5 lesiones aftosas con mucosa normal interlesional o áreas salteadas de lesiones de mayor tamaño	
<b>i-3</b>	Ileitis aftosa difusa sobre mucosa inflamada de forma difusa	Ileitis aftosa difusa sobre mucosa inflamada de forma difusa	
<b>i-4</b>	Inflamación difusa con presencia de úlceras grandes, nódulos y/o estenosis.	Inflamación difusa con presencia de úlceras grandes, nódulos y/o estenosis.	

\*Imágenes cedidas de la Unidad de Endoscopia Digestiva



### 2.3.2. Curación mucosa.

#### a. Importancia:

Desde hace algunos años ha cambiado el paradigma del tratamiento en la enfermedad inflamatoria intestinal y por tanto en la EC. Su objetivo actual no es mantener la remisión clínica sino alcanzar la curación mucosa, es decir, la ausencia de lesiones, para evitar el desarrollo de las complicaciones.

En múltiples estudios y en meta análisis recientes se ha mostrado cómo la curación mucosa se relaciona de forma significativa con un mejor pronóstico tanto a corto como a largo plazo en términos de remisión clínica mantenida y mejoría de la calidad de vida<sup>149</sup>, reducción de la necesidad de corticoides sistémicos, de ingresos hospitalarios y de intervenciones quirúrgicas<sup>150-152</sup>. Por ello a partir del año 2012 se convirtió en un objetivo primario no sólo en los ensayos clínicos sino también en la práctica clínica diaria, a pesar de la dificultad para valorarla.

Esta implicación directa de la curación mucosa en la práctica clínica diaria ha hecho que los índices endoscópicos, como método objetivo de la descripción de las lesiones endoscópicas, tengan cada vez mayor importancia.

#### b. Definición endoscópica de respuesta endoscópica, remisión endoscópica y curación mucosa:

Inicialmente la remisión endoscópica o curación mucosa se definía por características puramente descriptivas, la *International Organization For the Study of Inflammatory Bowel Disease* (IOIBD) la consideró como la ausencia de úlceras y erosiones, no obstante, esto es una definición muy simplificada debido a la afectación transmural de la enfermedad<sup>153</sup>. Además, no está validada ni permite cuantificar el empeoramiento o la mejoría de la afectación endoscópica.

Por ello, progresivamente, con el objetivo de evaluar de forma más objetiva la respuesta a los tratamientos se han ido introduciendo los índices de actividad endoscópicos (CDEIS o SES-CD), inicialmente en forma de ítems secundarios y en

los últimos años como ítems primarios para evaluar la eficacia de los tratamientos, fundamentalmente de las terapias biológicas<sup>154-156</sup>. El problema es que no disponemos de una definición validada de lo que se considera respuesta endoscópica, remisión endoscópica o curación mucosa.

Tras desarrollar el CDEIS, el mismo grupo, propuso una definición de curación mucosa como una puntuación del CDEIS <3, remisión endoscópica como una puntuación CDEIS < 6, y respuesta endoscópica como una reducción de al menos 5 puntos en el CDEIS<sup>157</sup>.

Tras este trabajo se han publicado otros estudios que han propuesto otras definiciones. Las definiciones de respuesta endoscópicas varían, desde una reducción del CDEIS entre 3-5 puntos hasta una reducción del CDEIS  $\geq$  50-75%<sup>155,156,158,159</sup>.

Ocurre lo mismo con la definición de remisión endoscópica; en los ensayos clínicos en los que se ha utilizado el CDEIS, se considera desde una remisión parcial, CDEIS < 6<sup>143,159</sup>, hasta la remisión completa con una puntuación que varía de 0 a  $\leq$  4<sup>143,155,159-161</sup>.

Sin embargo, no sólo se necesita una definición de lo que se considera respuesta o remisión endoscópicas, ya que lo que se pretende es alcanzar la **curación mucosa**. Todavía son muchos los estudios que utilizan el criterio morfológico propuesto por la *International Organization For the Study of Inflammatory Bowel Disease* como definición de curación mucosa. Entre ellos, el ACCENT I<sup>162</sup>, SONIC<sup>163</sup>, ENANCT-1<sup>164</sup> o incluso el EXTEND (*Extend the Safety and Efficacy of Adalimumab through Endoscopic Healing*)<sup>155</sup> y el MUSIC<sup>159</sup>, dos estudios pioneros en valorar la curación mucosa como objetivo primario. Posteriormente ya se han publicado estudios que incorporan estos índices endoscópicos para definir la curación mucosa, pero con definiciones variables que abarcan desde una puntuación CDEIS < 3<sup>165,166</sup> o una puntuación SES-CD < 5 o simplemente un cambio en la puntuación en ambos índices<sup>167</sup>. Sin embargo, los mejores resultados se siguen obteniendo con puntuaciones de ambos índices de 0<sup>168</sup> (Tabla 9).

Actualmente el abanico de lo que se considera curación mucosa es muy amplio, y la ausencia de úlceras se podría considerar como una definición parcial debido a la afectación transmural en la EC<sup>169</sup>. A pesar de ello, según la guía de la ECCO, es la definición que recomienda o bien una puntuación del CDEIS/SES-CD de 0.

Ante esta variabilidad en las definiciones y ausencia de validación de las mismas, recientemente Vuitton et al<sup>170</sup>, intentaron buscar una definición sobre remisión endoscópica mediante el método Delphi basado en una reunión de expertos. Se establecieron varias definiciones tanto de respuesta endoscópica como de remisión endoscópica para enfermedad luminal, utilizando el CDEIS y el SES-CD.

Como definición de respuesta endoscópica el mayor porcentaje de votaciones fue para la reducción del CDEIS o SES-CD mayor del 50%. Para remisión endoscópica SES-CD 0-2 fue la opción más votada.

Prácticamente no hay datos para definir curación en el contexto de recurrencia utilizando el índice de Rutgeerts. En la mayoría de los estudios se considera recurrencia endoscópica un  $i \geq 2$ <sup>131,171</sup>. Recientemente de Cruz et al publicaron un estudio en el que consideran respuesta y remisión endoscópica puntuaciones menores a 1<sup>136</sup>.

Es necesario diseñar estudios prospectivos dedicados exclusivamente para este fin y establecer una definición universal de lo que se considera curación mucosa y respuesta endoscópica, tanto para enfermedad luminal como en el contexto postquirúrgico.

Tabla 9. Resumen de algunas definiciones de respuesta endoscópica, remisión endoscópica y curación mucosa.

<b>Estudio</b>	<b>Respuesta endoscópica</b>	<b>Remisión endoscópica</b>	<b>Curación mucosa</b>
Mary et al. 2005 <sup>157</sup>	Reducción CDEIS>5	CDEIS <6	CDEIS <3
Van der Woule et al. 2010 <sup>158</sup>	Reducción CDEIS>3	----	----
Rutgeerts et al. 2006 (ACCENT 1) <sup>162</sup>	Reducción SES-CD		No úlceras
Colombel et al. 2010 <sup>163</sup>			No úlceras
Rutgeerts et al. 2012 (EXTEND) <sup>155</sup>	Reducción CDEIS>75% basal	CDEIS <4	No úlceras
Sandborn et al. 2012 <sup>154</sup>			No úlceras
Hebuterne et al. 2013 <sup>159</sup>	Reducción CDEIS>5	CDEIS <6 (<3 completa)	No úlceras
Ferrante et al. 2013 (SONIC) <sup>156</sup>	Reducción CDEIS o SES-CD >50% basal	--	No úlceras
Sakuraba et al. 2013 <sup>172</sup>	--	--	Completa: Reducción SES-CD >70% basal Parcial: reducción 20-70%
Vuitton et al. 2015 <sup>170</sup>	--	--	SES-CD 0-2
Sandborn et al. ENANCT-1 <sup>164</sup>	--	--	No úlceras
Noman et al. 2017 <sup>173</sup>	--	--	No úlceras

\*\*No disponemos datos sobre golimumab en EC.

## 2.4. Valoración de los Índices de Actividad. Función del Lector.

El objetivo de utilizar los índices endoscópicos es disponer de un mismo lenguaje endoscópico independientemente de quien realice la endoscopia, es decir, buscar una definición lo más objetiva posible y para ello se debe hacer énfasis en una de las características que deberían ser inherentes a los índices como es la fiabilidad o reproducibilidad.

### a. Reproducibilidad de los índices endoscópicos

La reproducibilidad es una propiedad fundamental para validar un índice en los ensayos clínicos. Es tan importante la variabilidad intraobservador como la variabilidad interobservador.

#### *Fiabilidad o reproducibilidad del índice CDEIS:*

De momento se han publicado 6 artículos que aportan datos sobre la fiabilidad del CDEIS, centrándose sobre todo en la concordancia interobservador. El primero de ellos, el estudio de validación, mostró una buena correlación en la valoración de la gravedad entre endoscopistas con un coeficiente de Pearson de 0,83 y también una elevada reproducibilidad interobservador entre 12 endoscopistas con una muestra de 75 vídeos, con una CCI 0,96  $p \leq 0,001$ <sup>142</sup>.

Recientemente dos estudios han confirmado esa buena reproducibilidad, en uno de ellos con un ICC intra e interobservador entre lectores centrales de 0,89 (IC 95% 0,86-0,93) y 0,71 (IC 95% 0,63-0,76)<sup>174</sup>. Y otro de ellos, entre gastroenterólogos especialistas en EC, con un ICC de 0,83 (IC 95% 0,54-0,99)<sup>175</sup>. Posteriormente en un análisis *posthoc* del estudio EXTEND, diseñado para evaluar la eficacia y seguridad del adalimumab para inducción y mantenimiento en EC ileocólica moderada-severa, se ofrecen datos sobre la concordancia interobservador entre lectores locales y centrales en diferentes semanas de tratamiento, confirmando esta buena reproducibilidad, en semana 0 [ CCI 0,78 (IC 95% 0,65-0,87)], semana 12 [ CCI 0,92 (IC 95% 0,86-0,95)] y semana 54 [ CCI 0,86 (IC 95% 0,73-0,93)]<sup>176</sup>. En otro estudio que evalúa la eficacia del certolizumab pegol en 89 pacientes con EC activa se obtienen datos similares aunque algo inferiores, el CDEIS basal presenta

una CCI 0,60 (IC 95% 0,39-0,75), mientras que en la semana 10 y en la 54 ofrece cifras superiores con un CCI 0,74 ( IC 95% 0,59-0,84) y CCI 0,81 (IC 95% 0,65-0,90) respectivamente<sup>159</sup>.

Aunque no se han definido puntos de corte relacionados con el pronóstico del paciente parece que el CDEIS es un índice que tiene una buena reproducibilidad tanto intra como interobservador según los datos disponibles actualmente.

#### *Fiabilidad y reproducibilidad del SES-CD.*

El SES-CD se definió posteriormente para facilitar la evaluación de la gravedad. De momento existen 4 estudios que evalúan la reproducibilidad de este índice. El primero de ellos, el estudio que definió el índice, con un CCI 0,98 (IC 95% 0,79-0,90)<sup>143</sup>, y posteriormente en 2014 el mismo autor reafirmó los datos en otro estudio en el que obtuvo un CCI 0,93 ( IC 95% 0,76-0,98)<sup>175</sup>. Al igual que el CDEIS, los últimos dos estudios publicados muestran una elevada reproducibilidad interobservador. En el análisis *posthoc* del EXTEND, en el que se compara la reproducibilidad entre lectores locales y centrales, se observa cómo se produce un incremento de la reproducibilidad conforme avanzan las semanas de tratamiento, en la semana 0 el CCI 0,77 (IC 95% 0,69-0,83), en la semana 12 CCI 0,86 (IC 95% 0,81-0,90) y en la semana 52 CCI 0,82 (IC 95% 0,73-0,88)<sup>176</sup>. Finalmente, el estudio de Khanna et al, en el que se evalúa el índice entre lectores centrales, se obtiene una reproducibilidad interobservador e intraobservador elevada con un CCI 0,83 (IC 95% 0,75-0,88) y de 0,91 (IC 95% 0,89-0,95) respectivamente <sup>174</sup>.

#### *Correlación entre los dos índices CDEIS-SES-CD.*

Varios estudios han determinado la buena correlación entre ambos índices; uno de ellos con un coeficiente de Spearman de 0,94<sup>177</sup> y otro con un CCI 0,92 (0,88-0,95)<sup>174</sup>.

Prácticamente no hay estudios publicados sobre la evaluación de la capacidad de respuesta de los índices endoscópicos. Recientemente se ha publicado un estudio que utilizando los videos del estudio EXTEND pretende valorar la capacidad de respuesta del CDEIS y SES-

CD, definiendo esa respuesta en función de dos parámetros que son el tratamiento administrado (adalimumab vs placebo) y el cambio en el *Patient Reported Outcomes PRO-2* (índice clínico basado en el dolor abdominal y frecuencia deposicional en los últimos 7 días). Según los resultados, ambos índices son capaces de detectar la respuesta endoscópica al tratamiento con ADA de forma moderada. Sin embargo, el SES-CD parece que es numéricamente superior al CDEIS, recomendando la utilización de este índice en los ensayos clínicos y en la práctica clínica diaria<sup>178</sup>.

#### *Fiabilidad del índice de Rutgeerts.*

No existen demasiados estudios que valoren la reproducibilidad del índice de Rutgeerts, sólo cuatro y con resultados similares.

En el estudio de Daperno et al<sup>175</sup>, el acuerdo interobservador entre 14 endoscopistas expertos fue moderado con un índice kappa de 0,57 (IC 95% 0,51-0,65), y bueno entre no expertos con un kappa de 0,67 (IC 95% 0,60-0,72).

En un estudio posterior en el que se incluyeron 4 lectores centrales, se obtuvo un grado de acuerdo intra e interobservador muy bueno y bueno con datos de CCI de 0,83 (0,75-0,90) y de 0,75 (0,63-0,85) respectivamente<sup>179</sup>.

En el contexto del estudio TOPPIC, los 5 lectores expertos realizaron un subanálisis para valorar el grado de acuerdo en el índice de Rutgeerts mediante fotografías e informes endoscópicos, que fue del 79%. Aunque el índice kappa global que se obtuvo fue de 0,50, cabe destacar que variaba de 0,72-0,88 dependiendo de la clasificación del índice. En este caso, si al considerar el punto de corte en  $i2$  (RPQ endoscópica según muchos estudios) el valor kappa para distinguir  $< i2$  vs  $\geq i2$  fue de 0,83<sup>180</sup>.

En un estudio reciente, 13 expertos realizaron la lectura de 39 colonoscopias (156 valoraciones) con el objetivo de valorar la variabilidad interobservador. El acuerdo interobservador global del índice fue moderado con kappa de 0,43 (IC 95% 0,33-0,52), sin embargo, la reproducibilidad se modificó según el punto de corte. La reproducibilidad en

distinguir i0-i1 vs i2-i4 fue moderada con valores de 0,47 (0,28-0,66), y se convirtió en buena en distinguir i0-i2 vs i3-i4 con valores kappa de 0,64 (IC 95% 0,42-0,85)<sup>181</sup>.

En otro estudio que incluyó a 58 gastroenterólogos su objetivo fue valorar la variabilidad en interpretar varios índices, entre ellos el índice de Rutgeerts<sup>182</sup>. El acuerdo interobservador del índice de forma global fue de 0,33 (IC 95% 0,28-0,38), sin embargo, ocurrió como en los estudios previos, con diferencias en función del punto de corte. Al considerar recurrencia  $\geq 2$ , el acuerdo fue de 0,44 (IC 95% 0,24-0,62), pero al determinar la recurrencia severa (i3-i4) entonces el grado de acuerdo ascendió a 0,54 (IC 95% 0,36-0,71). Los datos de los últimos dos estudios fueron muy similares y mostraron que efectivamente el grado de acuerdo global del índice de Rutgeerts es inadecuado, aunque si se divide en función de recurrencia grave o no, los valores mejoran discretamente.

Recientemente se ha publicado una revisión Cochrane sobre los dos índices endoscópicos más utilizados en la EC, el CDEIS y el SES-CD, evaluando la información disponible sobre distintas características. Aunque son muy utilizados con una gran fiabilidad, no se considera que estén completamente validados<sup>183</sup>. El índice de Rutgeerts queda fuera de esta revisión por lo que se desconoce si cumple todas las propiedades que debe cumplir un índice para su uso en ensayos clínicos y en práctica clínica diaria.

#### b. Lector central en la endoscopia de la EC.

Para realizar una valoración objetiva de la actividad inflamatoria no sólo se necesitan índices endoscópicos fiables, sino también un lector que sepa utilizar esa herramienta de la forma más objetiva posible.

Los índices endoscópicos tienen una variabilidad interobservador no despreciable, como hemos determinado previamente, además se ha observado cómo los lectores locales tienden a sobreestimar estos índices, con implicaciones importantes en la toma de decisiones terapéuticas del paciente. Debido a ello y con la intención de minimizar este sesgo teórico, en los últimos años se ha desarrollado la figura del revisor o lector central. El lector central o como lo ha denominado la FDA en una de sus guías, el *off-site reader*, es un concepto



relativamente nuevo. Lo definen como la evaluación de las imágenes o los videos endoscópicos fuera de los centros que están involucrados en el estudio y por personal que no tenga relación con los pacientes, investigadores o cualquier persona relacionada<sup>184</sup>. Se espera que al ser profesionales con más experiencia y más manejo de los índices que se utilizan en los estudios, acumulen menos errores en la lectura. Además, que sea ciego es algo indispensable para las agencias reguladoras. Hay guías de práctica clínica que regulan la importancia y el beneficio del *off-site reader* en otras especialidades, así como su diseño de lectura, que está ampliamente documentado<sup>184,185</sup>. Sin embargo, en el caso de los estudios endoscópicos y concretamente en la EII, la experiencia sobre esta nueva figura es muy limitada ya que es un concepto relativamente nuevo y, por tanto, no disponemos de guías.

Los primeros estudios que empezaron a tratar el tema de los *off-site reader* en la EII fueron los que se basaban en la CU. En uno de ellos que pretendía evaluar la eficacia de la mesalazina, la introducción de este lector conllevó una reducción del efecto placebo de 20% a 13,8% lo que produjo un incremento del efecto de la mesalazina de 9% a 15,2%, quedando patente la importancia de la objetividad en la evaluación de los pacientes para su inclusión en ensayos clínicos<sup>186</sup>.

En el caso de la EC los datos son más escasos, se estableció la figura del *off-site reader* tras el estudio MUSIC que evaluaba el certolizumab pegol para el tratamiento de la EC activa. En este estudio este lector ofreció una puntuación CDEIS inferior a la obtenida por los lectores locales, sobre todo en la endoscopia basal, a pesar de no encontrar una diferencia estadísticamente significativa<sup>159</sup>. En estudios posteriores como en el EXTEND, diseñado para valorar la eficacia del adalimumab en la EC luminal activa, asumió el beneficio de un lector central y se recurrió directamente al *off-site reader* para valorar la actividad endoscópica y la respuesta al tratamiento utilizando el CDEIS<sup>155</sup>.

En los últimos años disponemos poco a poco de más información sobre esta figura en la EC ya que lo prioritario es confirmar si realmente es necesaria esta figura en la endoscopia digestiva y si aporta los mismos beneficios reportados en otras especialidades en las que su uso está más extendido.

Se han publicado varios estudios sobre el grado de acuerdo de los índices SES-CD y CDEIS incluyendo la figura del *off-site reader*. En uno de ellos, evaluaron la concordancia

interobservador tanto del CDEIS como del SES-CD entre 4 lectores expertos<sup>174</sup>. Los datos obtenidos mostraron una concordancia interobservador tanto del CDEIS como del SES-CD entre lectores expertos casi perfecta con cifras de CCI 0,71 y 0,83 respectivamente, con una concordancia intraobservador de 0,89 y 0,91 respectivamente.

En otro estudio, Daperno et al<sup>175</sup> compararon la correlación de ambos índices entre un grupo de expertos en enfermedad inflamatoria y un grupo de gastroenterólogos generales, considerados no expertos. Los índices luminales, el CDEIS y SES-CD, obtuvieron una concordancia interobservador buena (CCI > 0,60). A pesar de ello, al comparar el grupo de expertos frente a los no expertos, la concordancia que obtuvieron los expertos fue claramente superior, concretamente el CDEIS CCI 0,83 (0,50-1,00) vs CCI 0,67 (0,36-0,97) y el SES-CD, CCI 0,93 (0,76-1,00) vs CCI 0,68 (0,35-0,97). Paradójicamente, el grado de acuerdo obtenido en el índice de Rutgeerts fue moderado en los expertos (kappa 0,57) y bueno en los no expertos (kappa 0,68).

Posteriormente, en un análisis *post hoc* del estudio EXTEND, se determinó el grado de acuerdo entre los lectores locales y un *off-site reader*, tanto de los índices SES-CD y CDEIS de forma global, como por segmentos intestinales. Se incluyeron 129 pacientes de 13 centros con el SES-CD y sólo 57 pacientes de 6 centros con el CDEIS. De forma global el estudio mostró una buena correlación en la puntuación final tanto de SES-CD como de CDEIS obtenida por los lectores centrales y los locales (CCI > 0,74); así como para valorar la reducción del índice en más del 75%, sobre todo en el CDEIS. Sin embargo, se observó que en algunos tramos intestinales como colon derecho la correlación no fue tan buena y que, en general, la puntuación del índice obtenida por los lectores locales fue superior a la obtenida por el lector central.

La mayoría de los estudios realizados en este contexto, como acabamos de ver, se basan en los índices de actividad endoscópica luminal. Los datos para valorar el grado de concordancia entre distintos lectores del índice de Rutgeerts en la recurrencia postquirúrgica son escasos.

En el estudio de Daperno, la concordancia del índice de Rutgeerts fue inferior a la obtenida con los índices de actividad luminal y con un resultado paradójico, con datos moderados entre expertos, kappa 0,57, y bueno entre los no expertos kappa 0,68<sup>175</sup>. En otro estudio

posterior que incluyó a 30 pacientes, Gecse et al<sup>179</sup>, evaluaron la variabilidad intra e inter observador del índice de Rutgeerts, del índice de Rutgeerts modificado (2a y 2b) y la valoración de gravedad mediante una escala visual. En esta ocasión los datos fueron más optimistas, ambos índices obtuvieron una buena concordancia intraobservador con datos de CCI 0,81 (0,72-0,88) y 0,83 (0,75-0,90) respectivamente y también interobservador aunque con valores inferiores 0,72 (0,60-0,83) y 0,75 (0,63-0,85).

## 2.5. Proceso de aprendizaje en lectura endoscópica.

### 2.5.1 Definición de la evaluación de competencias en la endoscopia digestiva.

Al igual que en otras disciplinas, en el ámbito sanitario resulta imprescindible adquirir las habilidades adecuadas para desempeñar las técnicas, entre las que se incluye la endoscopia digestiva.

En Medicina son necesarios los programas de aprendizaje que permitan confirmar que un médico ha adquirido los conocimientos y las habilidades adecuadas para desempeñar su trabajo. De hecho, esto se conoce desde los griegos y romanos, aunque no fue hasta 1260 cuando se estableció el primer sistema para certificar los conocimientos basados en un sistema escrito<sup>187</sup>.

Históricamente el esquema que se seguía en el campo del aprendizaje era el de *ver-hacer-progresar*, primero se observaban los procedimientos y posteriormente se repetían y perfeccionaban conforme aumentaban el número de exploraciones. Sin embargo, en los últimos años se ha producido un cambio sustancial en el campo del aprendizaje impulsado por dos movimientos, por una parte, asegurar la calidad de la endoscopia y por otra, la instauración del sistema de entrenamiento basado en competencias que disponga de un sistema objetivo capaz de evaluación.

En las últimas décadas la endoscopia digestiva ha ampliado su campo de trabajo, incrementando no sólo el número de exploraciones sino también la complejidad de éstas. A pesar de ello, la colonoscopia sigue siendo el procedimiento más realizado por su relación directa con la prevención y el diagnóstico precoz del cáncer colorrectal. Por ello, se han desarrollado, validado y monitorizado indicadores de calidad en la colonoscopia para asegurar que los programas de cribado sean de calidad.

Por tanto, resulta prioritario disponer de un buen sistema de aprendizaje de esta técnica que esté validado y que nos permita saber de una forma fiable que los endoscopistas han alcanzado los conocimientos y habilidades, es decir, la competencia, para realizar una colonoscopia de calidad<sup>188-190</sup>.

## 1. Definición de competencias en la Endoscopia Digestiva.

La definición de competencia ha cambiado en los últimos años. Durante mucho tiempo, según expertos o bien por consenso en diferentes sociedades médicas, se establecía según el número de exploraciones realizadas<sup>191,192</sup>. No obstante, se observó que el número de procedimientos era un aspecto más a tener en cuenta, pero no era el criterio más relevante en el aprendizaje.

Por este motivo, posteriormente se modificó la definición, y actualmente se considera competencia en la endoscopia como el mínimo nivel de habilidades, conocimiento y/o experiencia derivado tanto del aprendizaje como de la experiencia, que son necesarios para llevar a cabo un procedimiento de una forma segura y competente<sup>193</sup>.

Esta definición tan amplia, engloba tres esferas fundamentales<sup>194</sup> (Figura 4).

- Habilidades técnicas relacionadas con la técnica en sí. En el caso de la colonoscopia, cómo introducir el endoscopio para intubar el ciego (maniobras de entrada o rectificación), inspección adecuada de la mucosa y realización correcta de la terapéutica endoscópica.
- Habilidades cognitivas: saber aplicar la información obtenida en la práctica clínica, es decir, revisión de la historia del paciente (antecedentes médicos y quirúrgicos, tratamiento actualizado, etc) y la anatomía, explicar el procedimiento, preparación individualizada en cada caso, obtención del consentimiento informado, sedación administrada, hallazgos normales y reconocimiento de lesiones, saber realizar la terapéutica más adecuada en cada caso, el manejo de las complicaciones y redacción de un informe de forma adecuada con definiciones estandarizadas.
- Habilidades de integración: se trata de la fusión de las dos habilidades anteriores. En este punto juega un papel decisivo la experiencia y el comportamiento personal. Incluye múltiples aspectos como la toma de decisiones, el trabajo en equipo, el liderazgo, la profesionalidad o la educación entre otros.

Esta última es la parte más compleja, ya que es la que depende de la interacción de la propia técnica con los conocimientos que se han ido adquiriendo; pero

precisamente es la que va a permitir tomar una decisión en un equipo multidisciplinar y mejorar en la técnica cerrando el círculo del aprendizaje<sup>67</sup>.

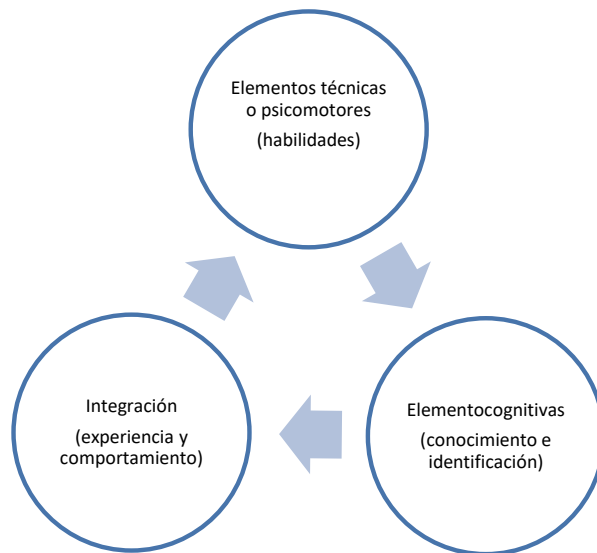


Figura 4. Esquema de la adquisición de competencias en la Endoscopia Digestiva. Adaptada del artículo de *Walsh et al*<sup>194</sup>.

## 2. Fases del aprendizaje (Figura 5).

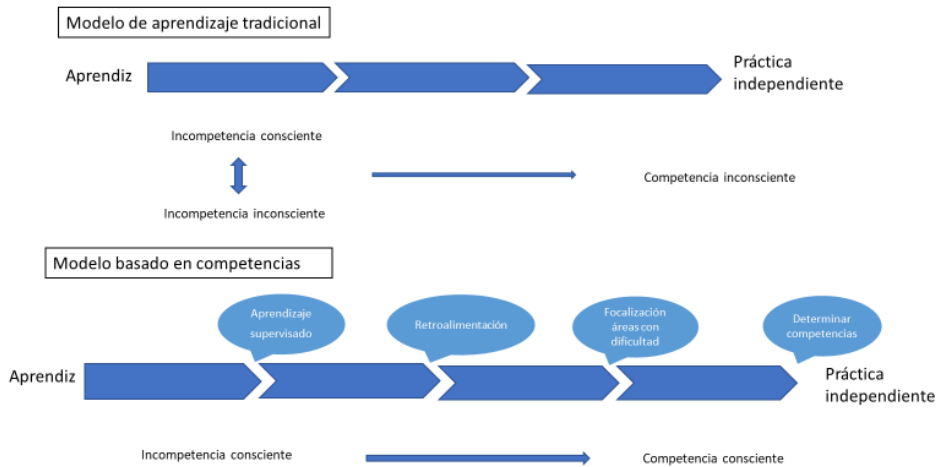


Figura 5. Diferencias entre sistema tradicional de aprendizaje y el basado en competencias. Adaptada del artículo de *Dubé et al* <sup>195</sup>.

Para alcanzar las competencias en la endoscopia digestiva es necesario adquirir estas tres habilidades y para ello es necesario seguir un proceso de aprendizaje que ha ido evolucionando con el tiempo.

En el modelo tradicional de aprendizaje el paso de incompetencia a competencia se produce tras un periodo básico de autoaprendizaje sin ningún tipo de tutorización ni de retroalimentación, con los errores acumulativos y los malos hábitos que esto supone.

Por ello, posteriormente se desarrolló un nuevo sistema de aprendizaje basado en competencias, en el que la retroalimentación con un tutor es un punto crucial del mismo. Gracias a este sistema, el alumno es capaz de pasar desde la incompetencia inconsciente a competencia inconsciente, permitiendo alcanzar las habilidades cognitivas (Figura 6).

La finalidad de este proceso es alcanzar ese estado de competencia consciente, que no se alcanza con el modelo tradicional, y que se obtiene gracias a la tutorización de todo el proceso de aprendizaje y focalización en las áreas de mayor dificultad para el alumno.



Figura 6. Fases del aprendizaje. Adaptado del artículo *Dubé et al* <sup>195</sup>.

### 3. Sistema de evaluación en el aprendizaje de competencias<sup>194</sup>.

Por supuesto, si el modelo de aprendizaje se modifica y se instaura un modelo basado en competencias, debemos disponer de un sistema de evaluación objetivo con unos ítems claros y unas herramientas sensibles y validadas que nos permitan saber si la evolución del proceso de aprendizaje es satisfactoria.

En primer lugar, los aspectos a evaluar en una técnica como la endoscopia digestiva, los podemos dividir en tres grandes bloques<sup>196</sup>.

- Bloque diagnóstico: se basa en obtener información sobre el conocimiento inicial de los alumnos, tanto a nivel teórico como práctico, así como los conceptos erróneos que tienen. Los métodos que se podrían utilizar son: cuestionarios con respuestas múltiples o establecer diferentes escenarios durante la endoscopia para tomar decisiones.
- Bloque formativo: centrado exclusivamente en el proceso de enseñanza. Se centra en los puntos más conflictivos detectados durante el bloque diagnóstico. Se les proporciona información relevante y algunos puntos clave para que los alumnos sean capaces de reflexionar sobre su rendimiento y puedan modificar su comportamiento para encauzar ese proceso de aprendizaje. En este apartado el *feedback* es fundamental porque es el que permite al alumno recapacitar y darse cuenta de sus carencias y/o errores en un tema en concreto, permitiendo solucionar dudas y



por tanto incrementando la motivación del aprendizaje. En este bloque también podrían utilizarse tests o simuladores.

- Bloque o evaluación de impacto: es cuando realmente se evalúa la competencia, es decir, es el momento de evaluar si el alumno ha alcanzado los objetivos del periodo de aprendizaje y por tanto ha tenido una progresión adecuada siendo capaz de realizar las exploraciones de forma autónoma. Esta evaluación se suele realizar mediante la observación directa durante la práctica clínica diaria.

Independientemente de la subdivisión que se haga, el aprendizaje es un proceso continuo en el que se imbrican los tres bloques, por lo tanto, la evaluación no debe ser en un momento determinado sino continua, ya que de esta forma se consigue el desarrollo personal y profesional para conseguir exploraciones cada vez más seguras y de más calidad.

En segundo lugar, es necesario disponer de herramientas para poder evaluar esos tres aspectos del aprendizaje, que sean fiables y estén validadas, permitiendo así conocer si se ha adquirido la habilidad y competencia oportuna. En el caso de la colonoscopia, existen distintas herramientas disponibles:

1. Número de procedimientos: históricamente era el método más utilizado para valorar la experiencia y capacidad de un endoscopista; habitualmente el valor se situaba entre 50-200, aunque podían ser muy variables. Con el tiempo se observó que, a pesar de que es necesario realizar un número determinado de exploraciones supervisadas antes de alcanzar las competencias adecuadas, no es el factor más importante para asegurar que se han alcanzado las habilidades técnicas y cognitivas<sup>193,197</sup>.
2. Datos métricos: tasa de intubación cecal, tolerancia a la prueba o tipo de preparación intestinal entre otras. A pesar de que son conceptos importantes, su uso como método de evaluación es controvertido ya que no incluye el *feed-back* con un experto para solventar dudas o carencias.

3. Evaluación mediante sistemas de simulación mecánicos o virtuales. Disponemos fundamentalmente de dos sistemas, los simuladores y los sistemas magnéticos. Los simuladores permiten al alumno desarrollar las habilidades técnicas y cognitivas en un ambiente seguro y con menos presión que en la práctica clínica diaria. En múltiples estudios se han observado como su uso en fases precoces reduce la curva de aprendizaje permitiendo alcanzar la intubación cecal en menos tiempo, con menor formación de asas y menor dolor para el paciente. Los sistemas magnéticos como el *Scopeguide*, a su vez, son muy útiles para detectar la formación de asas y aprender cómo reducirlas<sup>198,199</sup>.
4. Observación directa: probablemente sea la herramienta preferida debido a que permite la interacción entre profesor-alumno favoreciendo así el *feed-back*. Hasta ahora sólo en Reino Unido se ha incorporado de forma oficial esta herramienta en la evaluación del aprendizaje<sup>200,201</sup>. Hay múltiples herramientas descritas para evaluar las competencias en endoscopia digestiva baja. De ellas, 4 se encuentran validadas y se centran en evaluar las tres esferas que hemos nombrado previamente.
  - a. *Gastrointestinal Endoscopy Assessment Tool* (GiECAT) engloba las tres áreas de conocimiento. Se encuentra validado y presenta una elevada fiabilidad. No obstante, no existe una puntuación mínima a partir de la cual se considere que el alumno ha adquirido las competencias mínimas.
  - b. *Direct Observation of Procedural Skills* (DOPS) elaborado por el *Joint Advisory Group of Gastrointestinal Endoscopy* (JAG). Es un procedimiento directo de evaluación tanto diagnóstica como terapéutica, que está incluido en el sistema nacional de acreditación de competencias en la endoscopia digestiva del Reino Unido. Este documento consta de varios apartados, pre-procedimiento, procedimiento, post-procedimiento y *Endoscopic non-technical skills*. En cada uno de ellos, se incluyen ítems relacionados con la competencia y se puntúan de 1 a 4 (1 se considera que no ha alcanzado las habilidades y 4 que ha alcanzado la competencia adecuada). Para considerar al alumno apto, es necesario obtener valores superiores a 3 para cada ítem evaluado.

Actualmente se apuesta más por una escala de supervisión que por una puntuación concreta, ya que un valor numérico no es capaz de medir todos los

elementos que están relacionados con la adquisición de habilidades técnicas ni cognitivas que están relacionadas con un problema en concreto. Se han publicado estudios en los que demuestran una buena fiabilidad, una buena correlación con cuestionarios que determinan el conocimiento de los candidatos<sup>202</sup> y la posibilidad de discriminar entre endoscopistas experimentados de principiantes<sup>203</sup>.

- c. *Mayo Colonoscopy Skills Assessment Tool*, (MCSAT)<sup>204,205</sup>, se centra en la evaluación formativa y sobre todo en los aspectos intra-procedimiento. Es capaz de discriminar entre endoscopista experimentado de los que no lo son; no obstante, no se ha evaluado su fiabilidad de forma sistemática. La puntuación también es un score que grada de 1-4 al igual que el anterior y puntuaciones superiores a 3 son las que se consideran válidas.
- d. *ACE Colonoscopy Skill Assessment Tool*: desarrollado por la *American Society of Gastrointestinal Training Committee*, y basada en la anterior, se centra en la evaluación formativa. Se considera adecuada una puntuación superior a 3.5. No se ha estudiado su fiabilidad.

Todas estas herramientas de evaluación desarrolladas hace relativamente poco tiempo son parte del esfuerzo por mejorar la calidad asistencial evaluando los conocimientos de los endoscopistas. No obstante, se basan fundamentalmente en las habilidades técnicas.

### **2.5.2 Evaluación de habilidades cognitivas, valoración de lesiones. Curva de aprendizaje con material visual.**

En el apartado anterior queda reflejada la importancia de adquirir una formación adecuada en varias áreas, así como las herramientas disponibles hasta ahora para evaluarlas, que se centran mayoritariamente en las habilidades técnicas. Uno de esos ítems fundamentales en la endoscopia digestiva baja es la adquisición de las habilidades cognitivas, y en ella se incluye la interpretación de los hallazgos visualizados, ya sea durante la realización de la endoscopia en directo o bien un vídeo de ésta.

Actualmente no disponemos de información referente al método de enseñanza más adecuado para obtener esas habilidades cognitivas, entre las que se encuentran, la detección de lesiones sugestivas de EII, y concretamente de RPQ endoscópica. No obstante, podría ser un método de enseñanza muy similar al aprendizaje en la lectura de los vídeos de cápsula endoscópica.

La cápsula endoscópica es una técnica endoscópica que se basa en la valoración del tracto gastrointestinal mediante un vídeo realizado a partir de millones de fotografías. Múltiples artículos han demostrado el impacto positivo de los programas de aprendizaje en este contexto, no obstante, a pesar de ello, la mayoría de los métodos de aprendizaje se basan en cursos privados o patrocinados enmarcados dentro de las sociedades científicas, como ocurre en Inglaterra, Francia, Alemania, Italia, Portugal o España entre otras; así como también algún curso online. Actualmente no disponemos de una guía que nos permita saber cómo obtener esas habilidades cognitivas ni como evaluarlas.

Los cursos disponibles hasta ahora normalmente se dividen en lecciones teóricas, prácticas y sesiones interactivas donde se imparten los conceptos básicos de la cápsula endoscópica, el funcionamiento del software, las contraindicaciones y las indicaciones más frecuentes, la visualización de vídeos o fragmentos de estos para interpretar imágenes normales y de lesiones, así como la elaboración de informes utilizando la terminología CEST (*Capsule Endoscopy Structured Terminology*) para unificar los informes<sup>206</sup>. Todo esto lo desarrolla y conduce un grupo de expertos tanto nacionales como internacionales, realizando sesiones interactivas con casos clínicos reales junto con una discusión posterior explicando los hallazgos, las controversias y el diagnóstico final, así como solventando las dudas que puedan surgir<sup>207-210</sup>.

A pesar de la presencia de múltiples cursos de aprendizaje en este ámbito, no hay demasiada información sobre las competencias necesarias para ser considerado apto para realizar esta técnica. Las sociedades científicas todavía no se ponen de acuerdo con la duración de los cursos, el número de estudios supervisados y otros criterios y, de hecho, la mayoría se basa en opiniones de expertos<sup>211</sup>.

Se ha demostrado el impacto positivo en los programas de aprendizaje tanto presencial como por medio de programas informáticos mediante tests o vídeos<sup>212,213</sup>.

Ante la ausencia de recomendaciones establecidas para el aprendizaje de cápsula endoscópica, se ha publicado este mismo año un estudio sobre la experiencia de los directores de los cursos de aprendizaje de cápsula, a su vez considerados expertos, como una pequeña guía de cómo podría plantearse el aprendizaje de cápsula endoscópica, adaptado por supuesto a cada país<sup>214</sup>.

### **2.5.3. Competencias en endoscopia en la EII.**

En el caso concreto de la EII, extrapolando los datos de otros ámbitos de la endoscopia, también sería necesario un periodo de aprendizaje para poder interpretar de forma adecuada las lesiones identificadas y clasificarlas adecuadamente. Por ello, este aprendizaje de visualizar las lesiones e interpretarlas, podría ser similar al descrito en cápsula endoscópica.

Actualmente no existe un programa validado que permita obtener un certificado de competencias adquiridas en la endoscopia de la EII, aunque sí hay programas de aprendizaje para profundizar en este campo<sup>215</sup>.

En los últimos años, debido a la necesidad de utilizar índices para cuantificar la actividad inflamatoria, se han publicado algunos estudios que pretenden valorar la correlación de los índices o bien entre expertos o entre expertos y gastroenterólogos generales, así como si una intervención formativa puede suponer un beneficio. No obstante, no hay métodos formativos validados en este contexto.

Sólo hay dos estudios publicados en los que se valora el cambio de los índices endoscópicos tras un periodo formativo. El primero de ellos, publicado en el año 2016<sup>216</sup>, incluye a pacientes con EC y CU con actividad endoscópica. Un total de 237 gastroenterólogos sin experiencia en el uso de los índices endoscópicos que asistieron a varias reuniones científicas en un plazo de dos años consecutivos, realizaron una primera lectura determinando el índice de Mayo, el CDEIS, el SES-CD y el Índice de Rutgeerts en

unos 5-6 vídeos de cada categoría. Posteriormente, tras una discusión sobre las discrepancias o dudas en la lectura, realizaron una segunda lectura. En este estudio el CCI (IC 95%) para SES-CD y CDEIS fue de 0,77 (IC 95% 0,56-0,96) y 0,76 (IC 95% 0,54-0,96) respectivamente, y parece que en lo que se obtuvo más beneficio fue en el índice de Mayo que paso de un kappa 0,51 (IC 95% 0,48-0,55) a 0,76 (IC 95% 0,72-0,79) y en el Índice de Rutgeerts que pasó de un kappa de 0,45 (IC 95% 0,40-0,50) a 0,79 (IC 95% 0,70-0,80) ( $p < 0,0001$ ).

En el segundo estudio, se incluyeron a cinco gastroenterólogos no expertos y se evaluó la concordancia del SES-CD y del CDEIS utilizando 65 vídeos. Se observó una mejoría en el SES-CD tras la formación recibida con un CCI inicial de 0,78 (IC 95% 0,71-0,85) que pasó a 0,85 (IC 95% 0,79-0,89) mientras que el CDEIS no cambió prácticamente, pasando de 0,74 (IC 95% 0,65-0,81) a 0,73 IC 95% (0,65-0,81)<sup>217</sup>. En ambos estudios utilizaron una metodología similar para realizar el periodo formativo, que fue proporcionar los vídeos y tras la primera lectura realizar una sesión teórica con material teórico y vídeos en los que se ilustraron los aspectos más conflictivos de los índices para poder discutir los ítems más controvertidos.

A pesar de ello, hay muy pocos artículos publicados acerca de este tema por lo que se desconoce si es el sistema de formación más adecuado para poder instaurarlo de forma taxativa en la formación de los alumnos.

### **3. JUSTIFICACIÓN**

---





La necesidad de cirugía es frecuente en la EC, ya sea por las complicaciones derivadas de la enfermedad, como fístulas o estenosis, o bien por el fracaso del tratamiento médico. De hecho, la mayoría de los pacientes en algún momento de su evolución van a precisar al menos una intervención quirúrgica. No obstante, la cirugía no es curativa y la reaparición de la enfermedad en tramos no afectos previamente, en la anastomosis o vertiente ileal de la misma, es prácticamente la norma y es lo que se conoce como recurrencia post-quirúrgica.

Históricamente la valoración de la actividad inflamatoria se realizaba mediante los índices clínicos. Sin embargo, estos índices no se correlacionan con la actividad inflamatoria, debido a que la sintomatología que presentan los pacientes podía ser secundaria a la cirugía (síndrome de intestino corto o malabsorción) y no por la propia inflamación.

Por ello, en los últimos años la endoscopia y los índices endoscópicos tienen más protagonismo como método directo y objetivo de valorar la presencia y gravedad de la actividad inflamatoria. Para ello se dispone de un único índice endoscópico, que, aunque no esté validado, es el que se utiliza tanto en la práctica diaria como en los ensayos clínicos, el índice de Rutgeerts/índice de Rutgeerts modificado. Se conoce que la gravedad de las lesiones documentada de forma precoz (6-12 meses tras la cirugía) va a determinar un curso más agresivo o indolente de la enfermedad, con las implicaciones que esto conlleva en el tratamiento y seguimiento posterior.

A este respecto hay dos inconvenientes, por una parte, un único índice endoscópico no validado y subjetivo, el índice de Rutgeerts; y por otra parte, un lector, que es la persona que va a otorgar una puntuación determinada y ello va a depender tanto de sus conocimientos sobre la EC como su familiarización con el uso de los índices endoscópicos.

Actualmente, no existe un sistema de lectura en los ensayos clínicos (lectores locales, lector central, panel de expertos...) que pueda considerarse el patrón oro y que se pueda utilizar de referencia. La figura del lector central o el *offsite reader*, ya descrita en otros ámbitos de la medicina y con claras ventajas, fundamentalmente su objetividad, ha irrumpido en la endoscopia y sobre todo en la EC. A pesar de ello, actualmente se desconoce si realmente la figura de este lector es beneficiosa y si realmente es así, qué criterios debe cumplir.



## **4. HIPÓTESIS**

---



La introducción de un lector central en los paneles de lectura de imagen endoscópica para valorar la recurrencia post-quirúrgica en la enfermedad de Crohn podría reducir la variabilidad interobservador que es inherente al índice de Rutgeerts modificado, incrementando de esta forma el grado de acuerdo entre lectores expertos y no expertos.

Además, la formación en la lectura de los vídeos podría ser un factor que permitiese reducir la variabilidad interobservador mejorando el grado de acuerdo entre los lectores, incluso equiparando los lectores no expertos a los expertos.



## **5. OBJETIVOS**

---





**El objetivo general del estudio fue evaluar el grado de acuerdo intra e interobservador entre expertos clínicos en EII y expertos en endoscopias** en la clasificación del índice de Rutgeerts modificado en pacientes con EC intervenidos quirúrgicamente, y evaluar si la inclusión de un lector central reduce la variabilidad interobservador mejorando la reproducibilidad del índice. También se analizó el impacto de un programa formativo en la variabilidad interobservador entre lectores no expertos.

Este objetivo general se desglosó en los siguientes objetivos específicos:

1. Valorar la variabilidad intraobservador del índice de Rutgeerts modificado en el grupo de expertos en endoscopias digestivas y en el grupo de clínicos expertos en EII.
2. Valorar la variabilidad interobservador del índice de Rutgeerts modificado de los lectores expertos tanto de forma global como comparada con el lector central.
3. Valorar la variabilidad interobservador pre y post-formación del índice de Rutgeerts modificado de los lectores no expertos, globalmente y comparado con el lector central.



## **6. MATERIAL Y MÉTODO**

---



## **6.1. Diseño del estudio**

Se diseñó un estudio clínico observacional en el que un grupo de clínicos expertos en EII y endoscopistas expertos evaluaron videos de pacientes con EC intervenida y anastomosis ileocólica para determinar la recurrencia post-quirúrgica mediante el índice de Rutgeerts modificado.

El estudio fue aprobado por el Comité Ético de Investigación Clínica del Hospital Universitari i Politècnic La Fe el 24 de mayo de 2016, (Anexo 1).

## **6.2. Material y Métodos.**

Los vídeos incluidos en el trabajo pertenecen a ileocolonoscopias de pacientes con EC, en seguimiento por la Unidad Multidisciplinar de EII del Área de Enfermedades Digestivas del Hospital La Fe, realizadas según práctica clínica habitual en la Unidad de Endoscopias Digestivas en el periodo comprendido entre enero de 2013 y mayo de 2016.

### **6.2.1 Equipamiento endoscópico.**

Todas las exploraciones endoscópicas fueron realizadas, bajo sedación profunda controlada por un anestesista en dos de las cinco salas de la Unidad de Endoscopia Digestiva (Figura 7). Cada una de las salas estaba equipada con un brazo quirúrgico donde se encontraba ubicada la pantalla, el procesador EXERA II o EXERA III (Olympus) ( Figura 8), una bomba de lavado, otra bomba de CO2 y una fuente de diatermia.

Se utilizaron endoscopios de la marca Olympus, modelo CF-180AF y CF-HQ190. (Figura 9).



Figura 7. Sala de trabajo de la Unidad de Endoscopia Digestiva.



Figura 8. Procesador EXERA III. Procesador EXERA II.



Figura 9. Endoscopio Olympus CF-180AL.

Los vídeos endoscópicos fueron grabados mediante el programa utilizado en la Unidad de Endoscopias, ENDOBASE, sin datos identificativos del paciente. Se capturaron con una resolución de 768x576 pixels y se exportaron desde ENDOBASE a formato MPG2 con una compresión a calidad alta (bitrate de 4.9 Mbps). Posteriormente, tras obtener los vídeos, se editaron, para asegurar el anonimato y reducir el tiempo de visualizado, y se guardaron en formato *windows media video* (WMV).

### **6.2.2. Selección de los videos: criterios de inclusión y exclusión.**

Inicialmente se evaluaron para la inclusión 77 vídeos.

Los criterios de inclusión y exclusión del estudio fueron los siguientes:

- Criterios de inclusión:
  - Videos de pacientes con EC a los que se les haya realizado una resección ileocecal o hemicolectomía derecha y presenten una anastomosis ileocólica (termino-terminal, termino-lateral o laterolateral tanto isoperistáltica como anisoperistáltica) (Figura 10).
  - Consentimiento informado para la utilización de los vídeos en el estudio.
  - Intubación del neoíleon entre 10-15cm.
  - Grabación de al menos 1 minuto del neoíleon.

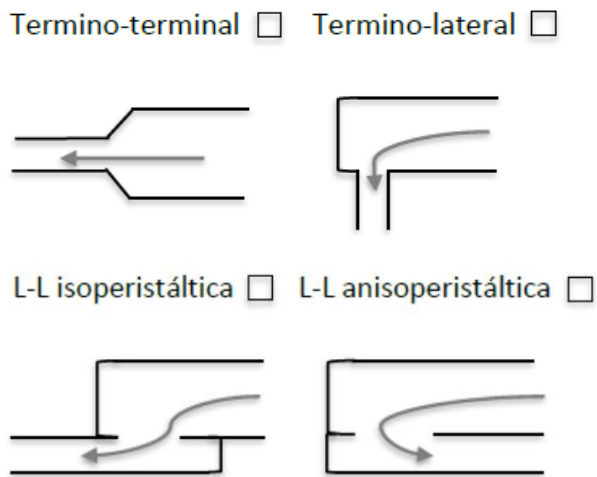


Figura 10. Tipos de anastomosis ileocólicas

- Criterios de exclusión:
  - Calidad del vídeo inadecuada según el coordinador del estudio.

De los 77 vídeos inicialmente disponibles se excluyeron 9 debido a calidad inadecuada, quedando finalmente 68 vídeos.



La clasificación inicial de los 68 vídeos según las categorías del índice de Rutgeerts modificado otorgadas por el lector local queda reflejada en la Tabla 10.

Tabla 10. Distribución inicial de los vídeos según el índice de Rutgeerts modificado otorgado por el lector local.

Valor	Frecuencia	Frecuencia Relativa
i0	11,00	0,16
i1	8,00	0,12
i2a	14,00	0,21
i2b	10,00	0,15
i3	3,00	0,04
i4	22,00	0,32

Debido a que se disponían de un mínimo de 8 vídeos de cada categoría ( excepto de i3 que sólo se disponían de 3 vídeos), se eligieron 8 vídeos de cada categoría mediante un sistema de muestreo aleatorio simple, excepto de la categoría i3, de la que se incluyeron los tres vídeos disponibles. Por lo tanto, se incluyeron un total de 55 vídeos para la lectura con la siguiente distribución:

- 8 vídeos originales de las categorías i0, i1, i2a,i2b, i4.
- 3 vídeos originales de la categoría i3.
- Además, se incluyeron 12 videos más, 2 repeticiones de un video de cada una de las categorías para calcular la variabilidad intraobservador.

Estos 55 vídeos elegidos se codificaron para que fueran anónimos y se editaron de forma que la duración máxima fuera de 3 minutos.

### 6.2.3. Plataforma de almacenamiento: Plataforma Owncloud La Fe.

Los vídeos seleccionados se almacenaron en un servidor del propio hospital con sistema de seguridad encriptado y utilizando el software libre de código abierto Owncloud (Figura 11) para el alojamiento de archivos en la nube. El acceso a la plataforma estaba limitado mediante usuario y contraseña.

Cada uno de los vídeos disponía de un enlace que es el que posteriormente se incluyó en los formularios para permitir el acceso directo y visualización de los vídeos.

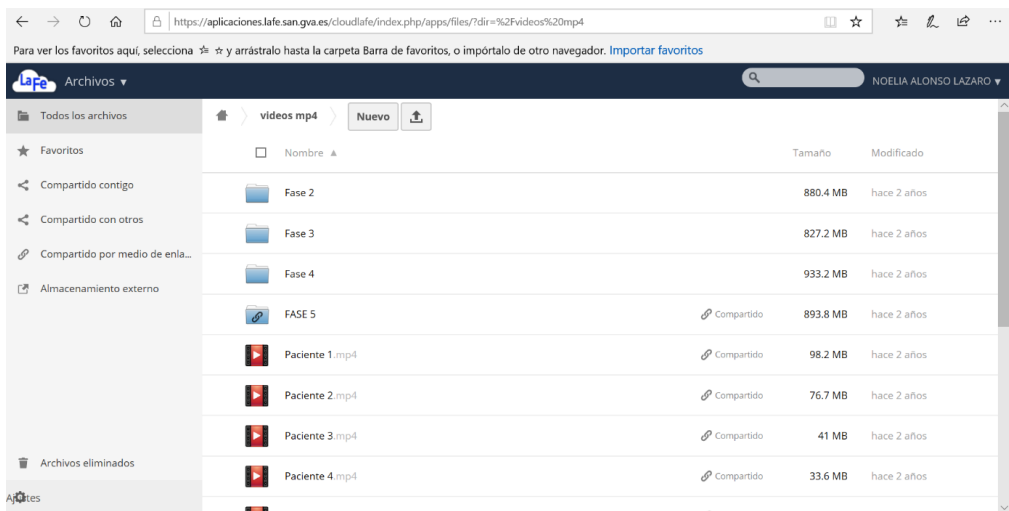


Figura 11. Plataforma owncloud.

### 6.2.4. Plataforma para la recogida de datos: Formularios Google.

Para facilitar la lectura de los vídeos y la recogida de datos de cada una de las evaluaciones por parte de los lectores se utilizó el recurso Formularios de Google como soporte de lectura, elaborando 5 formularios (Anexo 2).

El primer formulario incluyó, una breve introducción al estudio, la identificación de los lectores y el grupo al que pertenecían (Figura 12.1-12.2).

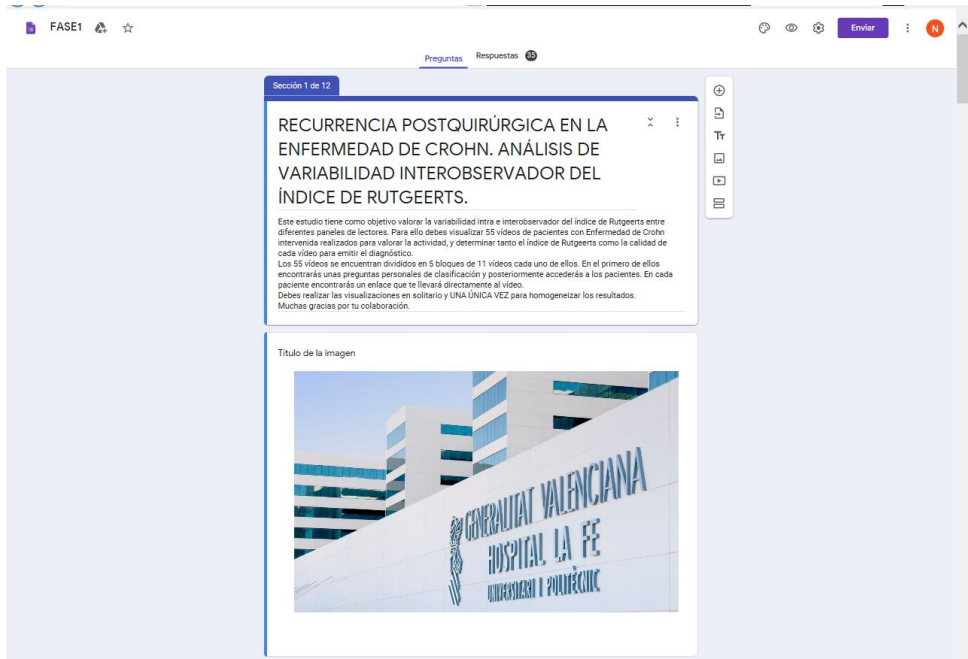


Figura 12.1. Primera parte del Formulario 1.

Nombre y apellidos \*

Tu respuesta

¿A que sección del hospital perteneces? \*

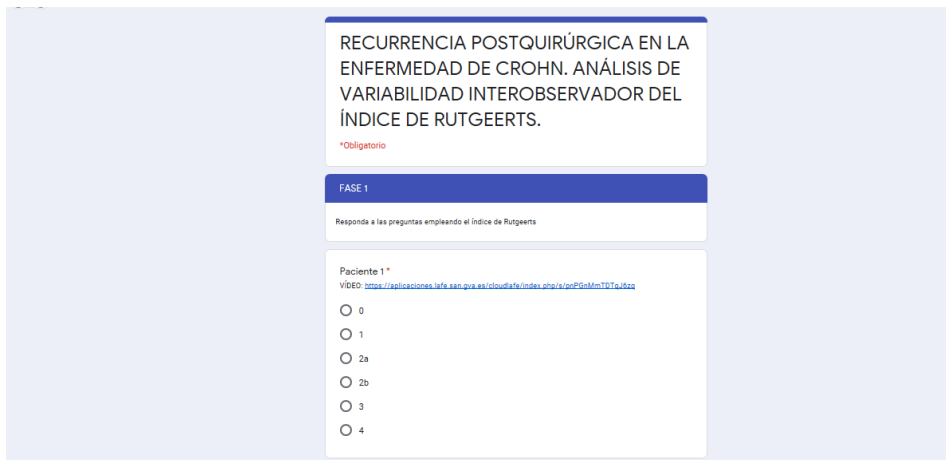
- Adjunto. Unidad de Endoscopia digestiva
- Adjunto. Unidad de Inflamatoria intestinal
- Residente de segundo año
- Residente de tercer año
- Residente de cuarto año

Siguiente

Figura 12.2. Segunda parte del Formulario 1.

Una vez cumplimentada esta primera parte, los lectores accedían a las páginas sucesivas. En cada formulario se incluyeron 11 vídeos y cada uno de ellos disponía de un enlace que los redireccionaba a la plataforma ownCloudLaFe para la visualización de cada uno de los vídeos.

En cada uno de los formularios, se debía cumplimentar de forma obligatoria el índice de Rutgeerts modificado (Figura 13). Una vez realizada la lectura de los 11 vídeos los resultados quedaban grabados en la plataforma, y posteriormente se podían descargar en formato Excel®.



RECURRENCIA POSTQUIRÚRGICA EN LA ENFERMEDAD DE CROHN. ANÁLISIS DE VARIABILIDAD INTEROBSERVADOR DEL ÍNDICE DE RUTGEERTS.

\*Obligatorio

FASE 1

Responda a las preguntas empleando el índice de Rutgeerts

Paciente 1 \*

VÍDEO <https://aplicaciones.lafe.san.gov.es/owncloud/fe/index.php/s/roP3iaM7DTaJbz2>

0

1

2a

2b

3

4

Figura 13. Página de formulario.

En segundo lugar, el coordinador del estudio envió un mail a todos los lectores, que previamente habían aceptado participar en el estudio, con un breve recordatorio del trabajo y algunas normas básicas para acceder a los vídeos y realizar la lectura, sin informar de la presencia de vídeos repetidos.

### 6.2.5. Lectores.

En total se incluyeron 18 lectores:

- Lectores no expertos: se consideraron a los médicos residentes en formación (MIR) con diferente grado de conocimiento tanto de la EC como de la endoscopia. Se incluyeron a 9 lectores: tres MIR de segundo año, otros tres MIR de tercer año y, por último, tres MIR de cuarto año.
- Lectores expertos: se incluyeron tres expertos en endoscopia digestiva y cinco expertos clínicos en EII. La elección de los expertos se realizó en base a su experiencia profesional. Los criterios que debían cumplir los expertos en endoscopias fueron, dedicación exclusiva en la endoscopia digestiva desde hacía más de 10 años y realizar más de 20 endoscopias de pacientes con EII al mes utilizando índices endoscópicos para valorar la actividad. Los criterios de los clínicos expertos en EII fueron, la dedicación preferente a la atención de pacientes con EII con el seguimiento de más de 40 pacientes/mes.
- Lector central: en nuestro trabajo se escogió a un experto en endoscopia digestiva, que además de cumplir los criterios expuestos previamente, debía tener experiencia concretamente en el índice de Rutgeerts modificado, y haber participado en ensayos clínicos de evaluación de recurrencia post-quirúrgica.

Todos los lectores disponían de un tiempo limitado (6 meses) para realizar la lectura, una vez pasado ese tiempo ya no era posible acceder a los vídeos. Además, debían tener en cuenta que una vez accedían a un formulario, debían realizar la lectura de todos los vídeos incluidos y enviarla posteriormente a la plataforma, ya que si la lectura no había sido completa, la información no quedaba grabada y debían repetir la lectura. Por ello se recomendó realizar la lectura en un ambiente tranquilo, con una conexión a internet estable y con un tiempo mínimo disponible de 60 minutos. También se recomendó evitar las últimas horas del día por el cansancio acumulado.

Los lectores expertos realizaron una única lectura. Sin embargo, en el caso de los lectores no expertos realizaron dos lecturas. Tras la primera lectura, tuvo lugar una intervención formativa de una tarde de duración en la sala de trabajo de la unidad de endoscopia digestiva del hospital La Fe.

Posteriormente, transcurridos dos meses, se les volvió a enviar el e-mail con los enlaces a los cinco formularios para que realizaran de nuevo la lectura, con las mismas recomendaciones que se habían dado previamente.

#### **6.2.6. Formación.**

Las sesiones formativas tuvieron lugar en la Unidad de Endoscopias del Hospital Universitari i Politènic La Fe y fueron impartidas por el coordinador del estudio y expertos en endoscopias y en EII. Consistieron en dos sesiones de dos horas de duración cada una de ellas.

- Sesión teórica en formato power point en la que se revisaron aspectos de la EII como la epidemiología, los criterios diagnósticos de Lennard-Jones, la importancia de la endoscopia en el diagnóstico y los índices de actividad endoscópica y su correlación con los índices de actividad clínica. Todo ello centrado en la recurrencia de la EC y en el papel de la endoscopia en la misma.
- Sesión práctica: se realizó una sesión con soporte audiovisual de 13 vídeos ileocolonoscopias no utilizados en la lectura. Se visualizó cada uno de los vídeos y en conjunto con el experto en endoscopias y el coordinador del estudio se discutieron los aspectos más controvertidos sobre los vídeos y cuál se consideraría el índice de Rutgeerts modificado más apropiado en cada caso.

#### **6.2.7. Variables del estudio.**

Se creó una base de datos anonimizada en formato Excel® donde se recogieron las variables demográficas (sexo, fecha de nacimiento, hábitos tóxicos, comorbilidades), las relacionadas con la enfermedad (fecha de diagnóstico, patrón de localización y de comportamiento de la enfermedad según la clasificación de Montreal, fecha de intervención, tipo de anastomosis, extensión de la resección y tratamiento en el momento de la inclusión) y con la endoscopia digestiva, fundamentalmente el índice de Rutgeerts

modificado para valorar la afectación inflamatoria, el tipo de tratamiento que llevaban durante la endoscopia y si se había modificado tras ésta ( Anexo 3).

#### **6.2.8. Evolución cronológica del estudio.**

En la siguiente figura (Figura 14), se puede visualizar el diagrama del estudio.

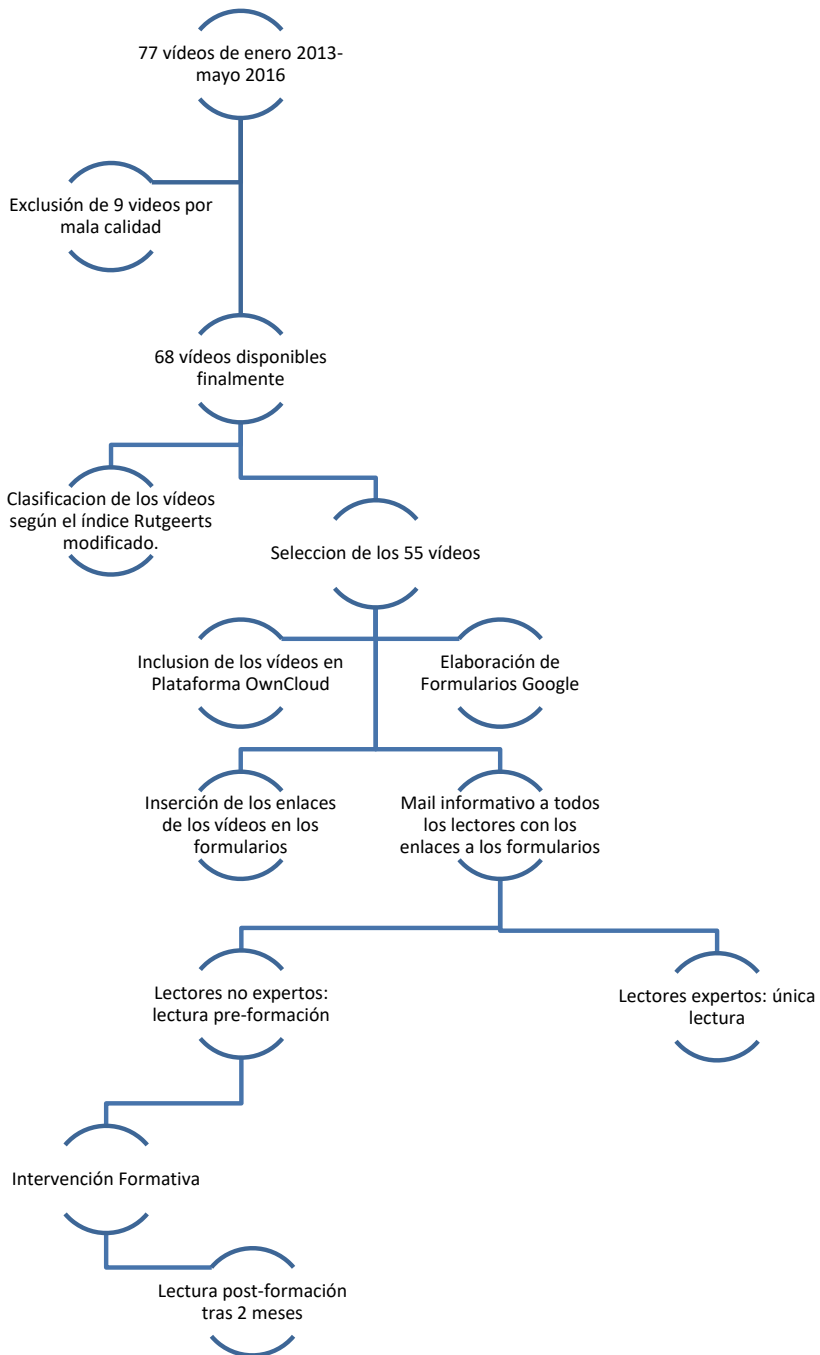


Figura 14. Diagrama cronológico del estudio.



## 7. ANÁLISIS ESTADÍSTICO

---



Los datos se presentan en forma de mediana y rango en caso de variables continuas/cuantitativas y mediante frecuencias y porcentajes en el caso de variables cualitativas/catóricas.

La variabilidad interobservador se definió como el grado de acuerdo entre los diferentes lectores durante la lectura de los 55 vídeos, mientras que la variabilidad intraobservador se definió como el grado de acuerdo de cada lector consigo mismo en los vídeos repetidos.

La variabilidad interobservador se calculó entre los dieciocho lectores. Para ello, se descargaron los archivos en Excel con las respuestas de todos los lectores y se estructuró la información recogida de forma que el índice de Rutgeerts modificado de cada vídeo se registró en filas y las respuestas de cada lector en columnas. Posteriormente, en otra hoja Excel, se reordenaron los datos para el cálculo de la variabilidad intraobservador, colocando los vídeos en filas y el índice de Rutgeerts modificado de cada visualización en columnas.

Se analizó la concordancia tanto del índice de Rutgeerts modificado como del índice agrupado en base a las siguientes definiciones de recurrencia/no recurrencia:

- 1ª agrupación: i0-i1 no recurrencia; i2-i3-i4 recurrencia.
- 2ª agrupación: i0-i1-i2a no recurrencia; i2b-i3-i4 recurrencia.
- 3ª agrupación: i0-i1-i2a-i2b no recurrencia grave; i3-i4 recurrencia grave.

El análisis de concordancia del índice de Rutgeerts modificado se realizó en base a tres parámetros, el porcentaje de acuerdo, el índice kappa de Fleiss y el coeficiente de concordancia de Kendall. Mientras que para el índice agrupado se realizó mediante el porcentaje de acuerdo y el índice kappa de Fleiss.

El índice kappa mide el grado de concordancia de evaluaciones nominales u ordinales realizadas por varios evaluadores. En este caso se utilizó el kappa de Fleiss ya que es una medida de acuerdo o reproducibilidad entre observaciones repetidas por más de dos lectores. Este estadístico tiene en cuenta si la puntuación es exactamente idéntica. En el caso de un vídeo puntuado como i0, penaliza de la misma forma si la puntuación otorgada es i1 o i4. Sus valores oscilan entre 0 y 1 y se interpretaron de la siguiente forma; < 0.20

concordancia pobre, 0,21 – 0,40 débil, 0,41 – 0,60 moderada, 0,61 – 0,80 buena y de 0,81 – 1.00 muy buena <sup>218</sup>.

El coeficiente de Kendall indica el grado de asociación de las evaluaciones ordinales realizadas por múltiples lectores, teniendo en cuenta la desviación de la misma, penalizando la puntuación en función de la desviación del índice basal. Su valor oscila entre 0 y 1 y se puede interpretar de forma similar al índice kappa o al CCI, valores superiores a 0,75 se consideraron adecuados <sup>219</sup>.

Los análisis estadísticos se realizaron utilizando el software Microsoft EXCEL 2010® (Microsoft, Redmon, WA USA) y el el software MINITAB® Statistical Software versión 18 (Enterprise Drive, Sate College, PA 16801 USA). Éste es un software para cálculos estadísticos que realiza un análisis de concordancia de atributos, es decir, clasificaciones subjetivas otorgadas por los lectores, en este caso el índice de Rutgeerts modificado o agrupado, para evaluar la concordancia entre las calificaciones realizadas por diferentes lectores.

La comparación del porcentaje de acuerdo del índice de Rutgeerts modificado entre varios lectores o en las tres agrupaciones se realizó mediante la prueba ANOVA. La comparación de dos grupos (expertos en EII *vs* expertos en endoscopias, expertos *vs* no expertos, expertos *vs* expertos incluyendo al lector central y pre-formación *vs* post-formación en los lectores no expertos) tanto del índice de Rutgeerts modificado como de las agrupaciones se realizó mediante el test de Wilcoxon.

Los p valores  $\leq 0,05$  se consideraron estadísticamente significativos.

En la Figura 15 se resumen los análisis de concordancia realizados entre los distintos grupos de lectores.

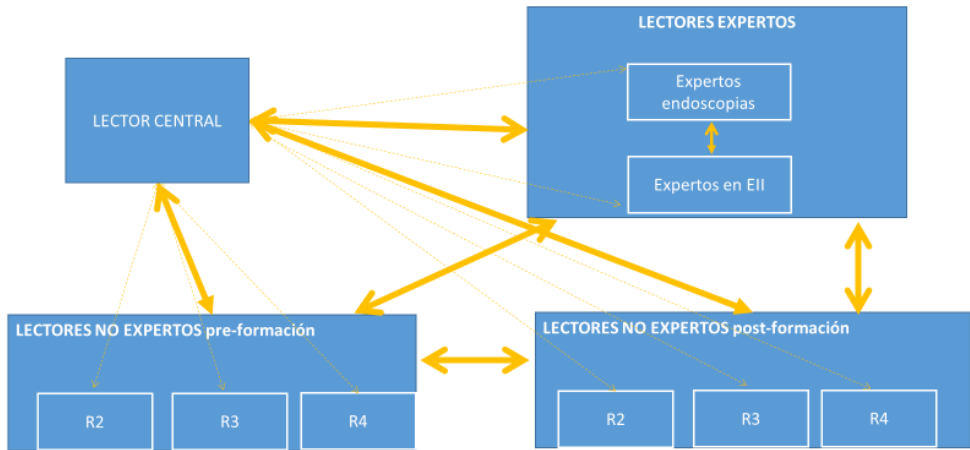


Figura 15. Resumen del análisis de concordancia realizado entre los lectores.

Cálculo del tamaño muestral:

El objetivo general del estudio fue evaluar el grado de acuerdo entre los diferentes lectores expertos utilizando el porcentaje de acuerdo y el índice kappa de Fleiss. El tamaño muestral se calculó estimando un porcentaje esperado de acuerdo para el grupo de expertos en endoscopias de 0,6 y de 0,3 para el grupo de expertos clínicos en EII, con una potencia del 85% y un nivel de confianza del 95%, obteniendo un tamaño muestral de 51 vídeos.

Finalmente se decidió ampliar a 55 videos (incremento del 7%) en previsión de problemas de lectura de los videos por una impresión subjetiva de mala calidad.



## **8. RESULTADOS**

---





## **1. Características basales de los vídeos y pacientes incluidos:**

Los vídeos incluidos pertenecían a 40 pacientes. El 55% eran mujeres con una mediana de edad de 40 años (rango, 18-73), no fumadoras (42,50%) y sin comorbilidades importantes (85%).

Un 67,50% se diagnosticó entre los 18-40 años, la localización más frecuente de la enfermedad fue ileocecal en un 47,50%, y el patrón de comportamiento predominante fue el estenosante en la mitad de los casos. Sólo un 32,50% de los pacientes asoció enfermedad perianal y la mediana de duración de la enfermedad desde el momento del diagnóstico fue de 14 años (rango 0-40).

La mayoría de las pacientes (67,50%) tenían sólo una resección quirúrgica (rango 1-4), cuya indicación fue la presencia de estenosis/cuadros suboclusivos en un 52,50% de los casos. Se realizó una resección ileocecal en un 62,50% y la anastomosis quirúrgica laterolateral fue la más frecuente (62,50%). Sólo se produjeron complicaciones quirúrgicas (intra y perioperatorias) en 3 pacientes (7,50%).

Tras la última resección quirúrgica 29 pacientes (72,50%) recibieron tratamiento preventivo de la recurrencia. El fármaco de elección en la mayoría de los casos (50,00%) fueron los inmunosupresores tiopurínicos.

Las características basales de los pacientes se recogen en las Tablas 11 y 12.

Tabla 11. Características demográficas de los pacientes incluidos.

<b>Datos demográficos</b>	<b>N=40 pacientes</b>
<b>Hombre/Mujer n/n (%/%)</b>	18/22 (45/55)
<b>Edad (mediana años, rango)</b>	40 (18-73)
<b>Hábito tabáquico n (%)</b>	
- <b>No Fumador</b>	17 (42,50)
- <b>Fumador</b>	12 (30,00)
- <b>Exfumador</b>	11 (27,50)
<b>Clasificación Montreal, n (%)</b>	
- <b>Edad al diagnóstico</b>	
<b>A1 (≤16 años)</b>	5 (12,50)
<b>A2 (17-40 años)</b>	27 (67,50)
<b>A3 (&gt;40 años)</b>	7 (17,50)
<b>No disponible</b>	1 (2,50)
- <b>Localización de la enfermedad colonoscopia</b>	
<b>L1 (ileal)</b>	21 (52,50)
<b>L3 (ileocólica)</b>	19 (47,50)
<b>L4 (afectación alta)</b>	3 (7,50)
- <b>Comportamiento en el momento de la colonoscopia</b>	
<b>B1 (inflamatorio)</b>	10 (25,00)
<b>B2 (estenosante)</b>	20 (50,00)
<b>B3 (penetrante)</b>	10 (25,00)
<b>Enfermedad Perianal</b>	13 (32,50)
<b>Duración de la enfermedad (años, rango)</b>	14 (0-40)

Tabla 12. Características quirúrgicas de los pacientes incluidos.

<b>Datos demográficos</b>	<b>N=40 pacientes</b>
<b>Indicación de cirugía, n (%)</b>	
- Actividad/Fracaso del tratamiento médico	15 (37,50)
- Estenosis/cuadros suboclusivos	21 (52,50)
- Fístulas	1 (2,50)
- No disponible	3 (7,50)
<b>Número de resecciones intestinales, mediana (rango)</b>	1 (1-4)
<b>Tipo de intervención, n (%)</b>	
- Resección ileocecal	25 (62,50)
- Hemicolectomía derecha	14 (35,00)
- No disponible	1 (2,50)
<b>Tipo de anastomosis, n (%)</b>	
- Laterolateral	25(62,50)
- Terminoterminal	6 (15,00)
- Terminolateral	1 (2,50)
- No disponible	8(20,00)
<b>Complicaciones, n (%)</b>	
- No	34 (85,00)
- Si	3 (7,50)
- No disponible	3 (7,50)
<b>Tratamiento prevención recurrencia tras la última cirugía, n (%)</b>	29 (72,50)
- Antibióticos	1 (2,50)
- Inmunosupresores tiopurínicos (IST)	20 (50,00)
- Biológicos (IFX/ADA)	1 (2,5) /4 (10)
- Tratamiento combinado: IST+ Biológicos	2 (4,50)
- Otros	1 (2,50)
- No disponible	27 (11,50)

## **2. Variabilidad intraobservador**

La variabilidad intraobservador se calculó en base a los vídeos repetidos durante la lectura.

El acuerdo intraobservador en el grupo de endoscopias fue del 100%, obteniendo un coeficiente de concordancia de Kendall de 1,00 y un índice kappa también de 1,00.

En el grupo de expertos en EII, tres de los cinco lectores, obtuvieron un grado de acuerdo del 100% con un índice kappa y un coeficiente de Kendall de 1,00. Otro lector obtuvo un porcentaje de acuerdo del 83,33% con un índice kappa de 0,75 y un coeficiente de Kendall de 0,69, y el último lector presentó un acuerdo del 66,67% con un índice kappa de 0,56 y un coeficiente de Kendall de 0,87.

El acuerdo intraobservador del lector central fue del 50,00 % con un IC 95% (11,81-88,19), con un índice kappa de 0,53 y un coeficiente de Kendall de 0,64.

La información sobre la variabilidad intraobservador se adjunta en el Anexo 4.

## **3. Variabilidad interobservador.**

Los vídeos fueron evaluados por dos grupos de expertos y por un lector central.

En el **grupo de endoscopia digestiva** el grado de acuerdo entre los lectores fue del 41,82% con un IC 95% (28,65-56,00) y un índice de kappa global de 0,47.

En el **grupo de clínicos expertos en EII** el grado de acuerdo fue del 14,55% con un IC 95% (6,50-26,66) y un índice de kappa de 0,32.

La diferencia del porcentaje de acuerdo entre ambos grupos fue significativa con una p 0,0357.

Al considerar a todos los **lectores como un único grupo** sin establecer diferencias entre endoscopistas y clínicos expertos en EII, el grado de acuerdo global fue del 7,27% con un IC 95% (2,00-18,00) y un kappa de 0,34.

La concordancia según las categorías del índice de Rutgeerts modificado queda reflejada en la Tabla 13.

Tabla 13. Concordancia interobservador por grupos de expertos y globalmente.

	<b>Endoscopistas</b>		<b>Experto clínico en EII</b>		<b>Global</b>	
<b>Respuesta</b>	kappa±DE*	p	kappa±DE*	P	kappa±DE*	p
<b>i0</b>	0,81±0,08	0,00	0,71±0,04	0,00	0,75±0,02	0,00
<b>i1</b>	0,39±0,08	0,00	0,25±0,04	0,00	0,29±0,03	0,00
<b>i2a</b>	0,34±0,008	0,00	0,23±0,04	0,00	0,27±0,03	0,00
<b>i2b</b>	0,49±0,008	0,00	0,07±0,04	0,05	0,14±0,03	0,00
<b>i3</b>	0,00	0,59	0,17±0,04	0,00	0,12±0,03	0,00
<b>i4</b>	0,31±0,008	0,00	0,40±0,04	0,00	0,37±0,03	0,00
<b>Global</b>	0,47±0,008	0,00	0,32±0,02	0,00	0,34±0,01	0,00
<b>Kendall</b>	0,82	0,00	0,77	0,00	0,74	0,00
<b>% acuerdo</b>	41,82		14,55		7,27	0,03
<b>IC 95%</b>	(28,65-55,89)		(6,50-26,66)		(2,02-17,59)	

\*DE: Desviación estándar

#### 4 . Agrupación del índice de Rutgeerts modificado en categorías:

Tras realizar las agrupaciones, el grado de acuerdo y el índice kappa obtenido tanto de forma global como por grupo de expertos queda reflejado en la Tabla 14.

Tabla 14. Variabilidad global y por expertos según las agrupaciones del índice de Rutgeerts modificado.

	<b>Global</b>	<b>Endoscopistas</b>	<b>EII</b>	<b>p</b>
<b><i>i&lt;2 vs ≥2</i></b>				
<i>% acuerdo</i>	50,91 (37,07-64,65)	72,73 (59,00-83,86)	58,18 (44,11-71,35)	0,14
<i>kappa±DE*</i>	0,56±0,03	0,61±0,08	0,55±0,04	
<b><i>i≤2a vs ≥2b</i></b>				
<i>% acuerdo</i>	36,36 (24,00-50,40)	72,73 (59,00-83,86)	47,27 (33,65-61,2)	0,07
<i>Kappa±DE*</i>	0,49±0,03	0,61±0,08	0,45±0,04	
<b><i>i&lt;3 vs ≥3</i></b>				
<i>% acuerdo</i>	43,64 (30,00-58,00)	78,18 (65,00-88,19)	60,00 (45,91-72,98)	0,14
<i>Kappa±DE*</i>	0,43±0,03	0,29±0,07	0,59±0,04	

\*DE: desviación estándar

- Al considerar RPQ endoscópica un valor  $\geq 2$ , el porcentaje de acuerdo global entre los lectores fue del 51% con un kappa de 0,56. Si se analizan los datos por separado según el grupo al que pertenecen los lectores, el porcentaje de acuerdo global entre los lectores expertos en endoscopia fue del 72,73%, con un kappa de 0,61 y en el caso de expertos en EII, 58,18% con un kappa de 0,55.

- Si se considera RPQ endoscópica  $\geq i2a$ , el porcentaje de acuerdo global entre todos los lectores fue del 36,36% con un kappa de 0,50. En el subgrupo de lectores endoscopistas fue del 72,73% con un kappa de 0,6 y en clínicos expertos en EII del 47,27% con kappa 0,45.
- Si se considera RPQ endoscópica grave un valor  $\geq i3$ , el porcentaje global de acuerdo obtenido fue de 43,64% con un kappa de 0,43. En el grupo de endoscopistas el porcentaje de acuerdo fue del 78% vs 60% el grupo de clínicos expertos en EII con un índice kappa 0,29 vs 0,59, respectivamente.

### **5. Lector central.**

En este apartado se evalúa la concordancia entre el grupo de expertos y el lector central.

Tras la lectura de los 55 vídeos, la Figura 16 muestra la puntuación calculado por todos los lectores expertos.

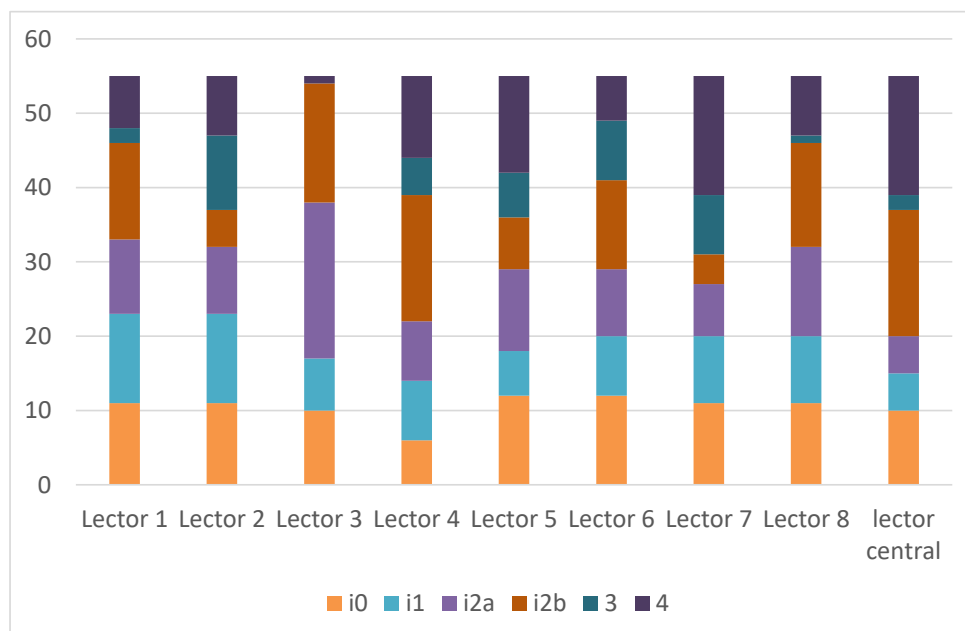


Figura 16. Puntuación de los vídeos según el índice de Rutgeerts modificado calculado por todos los lectores.

- Índice de Rutgeerts modificado:

Si se consideran a los **lectores de forma global** como grupo y se compara con el lector central el porcentaje de acuerdo fue del 7,27% con un kappa de 0,42 y un coeficiente de Kendall de 0,70.

El **grupo de expertos en endoscopias** comparándolo con el lector central, presentó un acuerdo de 29,09% con un IC 95% (17,60-42,09), índice de kappa de 0,35 y un coeficiente de Kendall de 0,69. Mientras que en el **grupo de expertos en EII**, el grado de acuerdo comparado con el lector central fue de 14,55% con un IC 95% (6,50-26,66), un kappa 0,45 y un coeficiente de Kendall de 0,70. La diferencia entre el grupo de expertos en endoscopias vs en EII no alcanzó la significación estadística,  $p = 0,14$ . (Tabla 15).

Tabla 15. Variabilidad comparada con el lector central del grupo de expertos.

	<b>Endoscopias</b>	<b>EII</b>
<b>% acuerdo</b>	29,09 (17,60-42,09)	14,55 (6,50-26,66)
<b>kappa±DE</b>	0,35±0,04	0,45±0,03
<b>Coficiente Kendall</b>	0,69	0,70

±DE: Desviación estándar

$P = 0,1429$

En la Figura 17 se muestra la concordancia analizada con el kappa tanto del índice de Rutgeerts modificado de forma global como de las categorías.



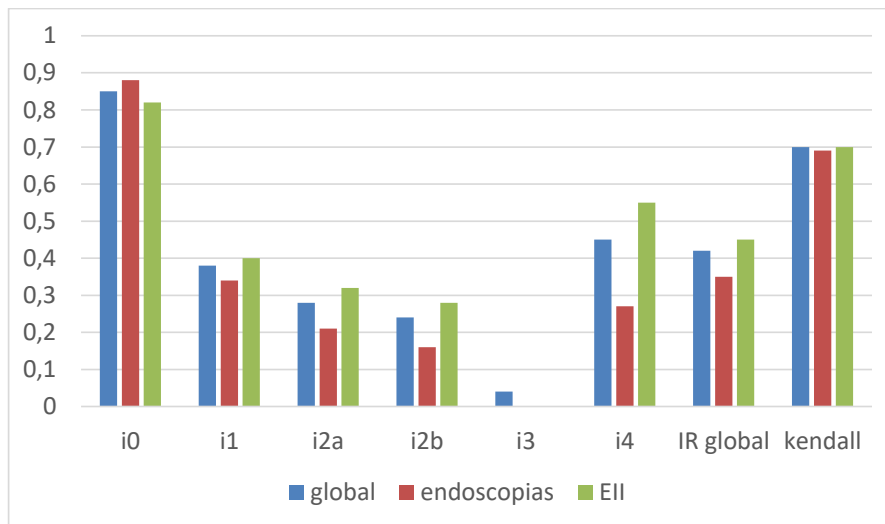


Figura 17. Concordancia según el valor kappa de los lectores comparado con el lector central.

También se comparó la lectura realizada por el grupo de expertos con la lectura de los expertos comparado con el lector central. Los lectores expertos obtuvieron un acuerdo del 7,27% con un IC 95% (2,02-17,59), y un índice kappa de 0,34. Sin embargo, al comparar su lectura con el lector central, el grado de acuerdo fue 7,27 (2,02-17,59) con un índice kappa de 0,42, tal y como se muestra en la Tabla 16, con una  $p = 0,832$ .

Tabla 16. Variabilidad del grupo de expertos comparada con el lector central.

	% acuerdo (IC 95%)	kappa±DE	Kendall
<b>Expertos grupo global</b>	7,27 (2,02-17,59)	0,34±0,01	0,74
<b>Expertos comparados con lector central</b>	7,27 (2,02-17,59)	0,42±0,02	0,70

±DE: Desviación estándar

$P=0,832$

- Agrupaciones del índice de Rutgeerts modificado:

El porcentaje de acuerdo de cada lector comparado con el lector central en cada una de las tres agrupaciones propuestas se muestra en la Tabla 17.

Tabla 17. Porcentaje de acuerdo de los lectores expertos comparado con el lector central.

	Lector 1	Lector 2	Lector 3	Lector 4	Lector 5	Lector 6	Lector 7	Lector 8	Media acuerdo IC 95%
<b><math>i &lt; 2</math> vs <math>\geq 2</math></b>	81,82	81,82	92,73	87,27	94,55	90,91	80,00	80,00	86,13 (81,90-90,4)
<b><math>i \leq 2</math> vs <math>\geq 2</math>b</b>	72,73	67,27	67,27	81,82	76,36	76,36	72,73	74,55	73,60 (69,4-77,90)
<b><math>i &lt; 3</math> vs <math>\geq 3</math></b>	80,00	89,09	69,09	92,73	83,64	89,09	81,82	83,64	83,60 (79,4-87,90)

P = 0,0012

Se calculó el grado de concordancia para las tres agrupaciones según el porcentaje de acuerdo y el índice kappa de forma global para todos los lectores así como según el grupo de expertos ( Tabla 18).

Tabla 18. Concordancia del índice de Rutgeerts agrupado en el grupo de expertos global y por categorías.

<b>Índice Rutgeerts</b>	<b>Global</b>	<b>Endoscopistas</b>	<b>EII</b>	<b>p</b>
<b><math>i &lt; 2</math> vs <math>i \geq 2</math></b>				
% acuerdo	50,91	69,09	58,18	0,39
IC 95%	(37,07-64,95)	(55,19-80,86)	(44,11-71,35)	
kappa $\pm$ DE*	0,68 $\pm$ 0,05	0,65 $\pm$ 0,08	0,69 $\pm$ 0,06	
<b><math>i \leq 2a</math> vs <math>\geq 2b</math></b>				
% acuerdo	36,36	56,36	47,27	0,25
IC 95%	(24,00-50,40)	(42,32-69,70)	(34,00-61,02)	
kappa $\pm$ DE*	0,47 $\pm$ 0,05	0,43 $\pm$ 0,08	0,49 $\pm$ 0,06	
<b><math>i &lt; 3</math> vs <math>i \geq 3</math></b>				
% acuerdo	43,64	65,45	60,00	0,39
IC 95%	(30,00-58,00)	(51,42-77,76)	(45,91-72,98)	
kappa $\pm$ DE*	0,56 $\pm$ 0,04	0,31 $\pm$ 0,07	0,71 $\pm$ 0,06	

\*DE: desviación estándar

También se evaluó la concordancia entre la lectura de los expertos como grupo global y la lectura de los expertos comparada con el lector central, como queda reflejado en la Tabla 19.

Tabla 19. Concordancia según varias agrupaciones del índice de Rutgeerts modificado en el grupo de expertos global, comparado con el grupo de expertos con el lector central.

<b>Índice de Rutgeerts</b>	<b>Expertos grupo global</b>	<b>Expertos comparados con lector central</b>	<b>p</b>
<b><math>i &lt; 2</math> vs <math>i \geq 2</math></b>			
% acuerdo IC 95%	50,91 (37,07-64,65)	50,91 (37,07-64,65)	0,96
kappa±DE*	0,56±0,03	0,68±0,05	
<b><math>i \leq 2a</math> vs <math>\geq 2b</math></b>			
% acuerdo IC 95%	36,36 (24,00-50,40)	36,36 (24,00-50,40)	0,96
kappa±DE*	0,49±0,03	0,47±0,05	
<b><math>i &lt; 3</math> vs <math>i \geq 3</math></b>			
% acuerdo IC 95%	43,64 (30,00-58,00)	43,64 (30,00-58,00)	0,87
kappa±DE*	0,43±0,03	0,56±0,05	

\*DE: desviación estándar

## **6. Cambio del lector central:**

Se realizó el mismo análisis cambiando el lector central con el objetivo de valorar si realmente la elección del lector central fue la adecuada. En vez de escoger un experto en endoscopias, se escogió, por una parte, un experto en EII y, por otra parte, mediante un sistema de muestreo aleatorio, otro lector dentro del grupo global de expertos. Se calculó el porcentaje de acuerdo de todos los lectores con cada uno de los expertos (Figura 18).

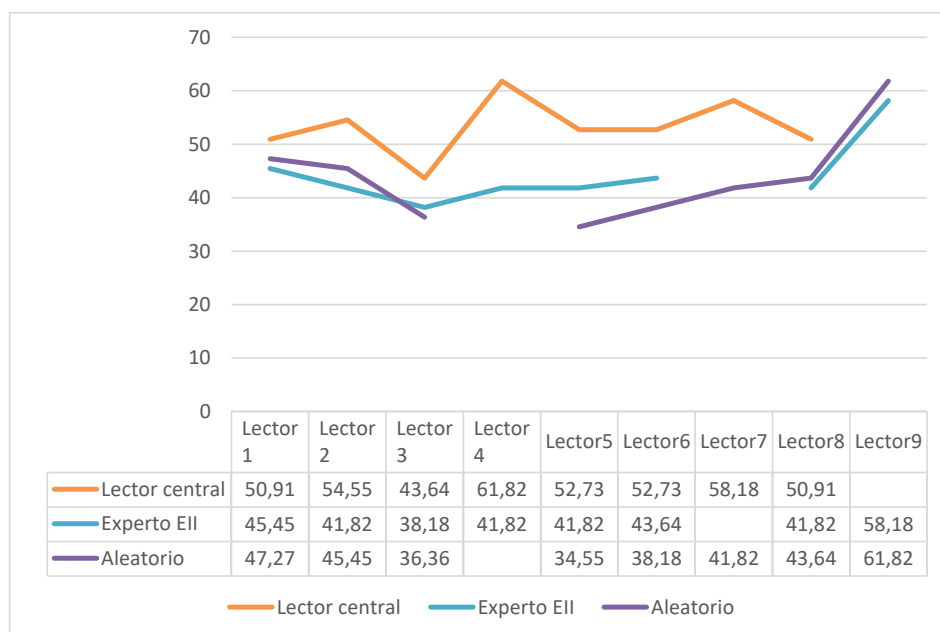


Figura 18. Porcentaje de acuerdo de los lectores expertos comparado con los lectores centrales.

La media de acuerdo de todos los lectores con el lector central inicial, experto en endoscopias, fue de 53,18% (IC 95% 48,17-58,19) vs 44,09% (IC 95% 39,08-49,10) con el lector central experto en EII vs 43.63% (IC 95% 38,62-48,66) con lector central escogido aleatoriamente, con una diferencia estadística ( $p = 0,0168$ ) (Figura 19).

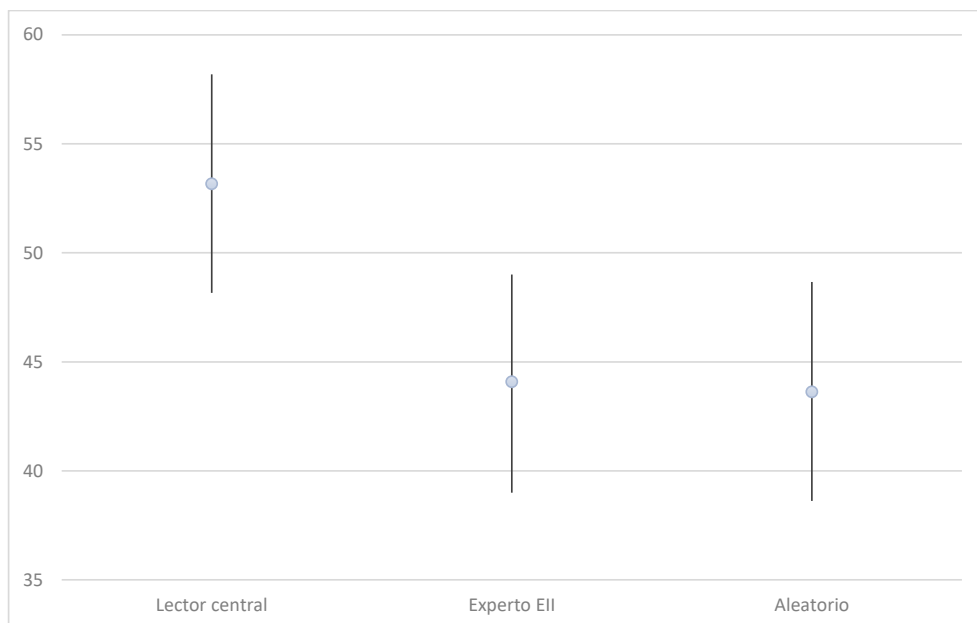


Figura 19 Porcentaje de acuerdo de los lectores comparado con los diferentes lectores centrales.

Posteriormente se realizaron los mismos análisis sustituyendo al lector central experto en endoscopias por el lector experto en EII, obteniendo los siguientes resultados.

- La variabilidad interobservador global fue del 11% (IC 95% 4,00-22,25) con un kappa de 0,37 y un coeficiente de concordancia de Kendall de 0,77. Comparando todos los lectores como grupo global con el lector central, el porcentaje de acuerdo fue 7% (IC 95% 2,00-18,00) con kappa de 0,32 y coeficiente de correlación de 0,57.
- Los expertos en endoscopias, en comparación con el lector central, obtuvieron un porcentaje de acuerdos de 20% (IC 95% 10,4-33,00), con un kappa de 0,34, y un coeficiente de Kendall de 0,58.

- Los expertos en EII, en comparación con el lector cental, obtuvieron un porcentaje de acuerdos de 14.55% (IC 95% 6,50-26,70) con un kappa de 0,29 y Kendall 0,57.

Los resultados se detallan en el Anexo 5.

## 7. Actividad Formativa:

La puntuación que obtuvieron los lectores no expertos pre-formación y post-formación se resume en las Figuras 20 y 21.

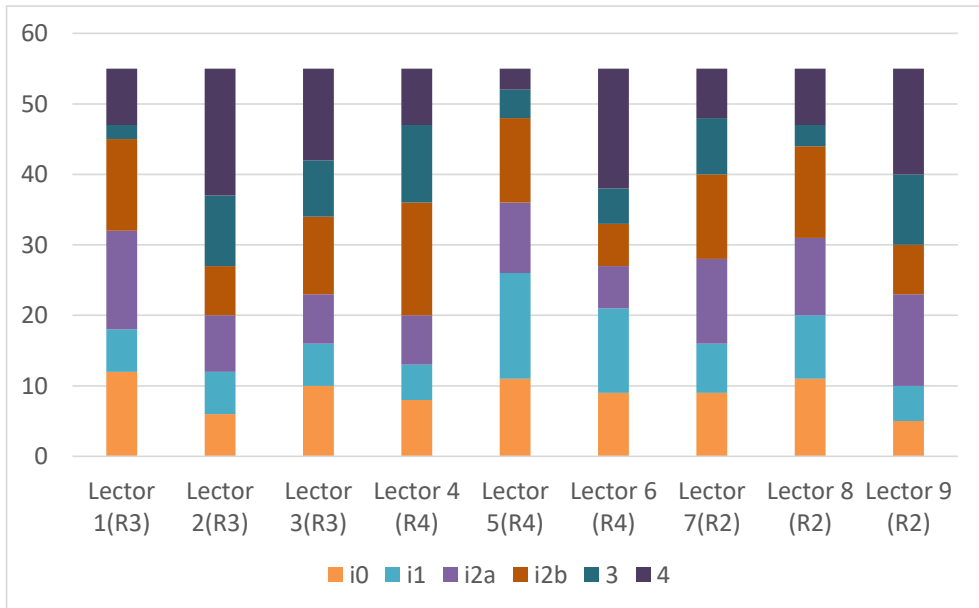


Figura 20. Distribución de la puntuación del índice de Rutgeerts modificado de los lectores no expertos pre-formación.



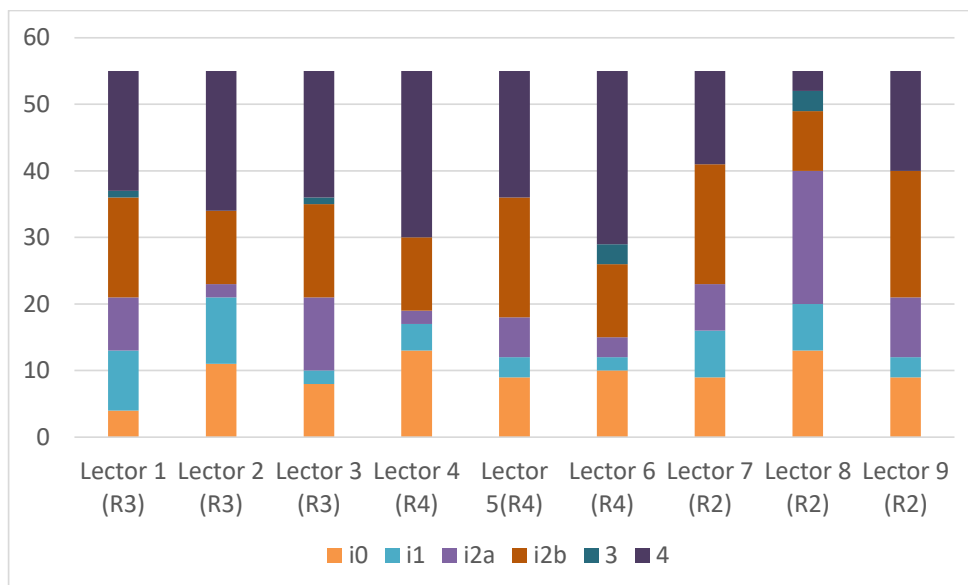


Figura 21. Distribución de la puntuación del índice de Rutgeerts modificado de los lectores no expertos post-formación.

Variabilidad interobservador:

En la lectura pre-formación, el porcentaje de acuerdos fue 3,64% con un IC 95% (0,44-12,53), un índice kappa global 0,29 y un coeficiente de Kendall 0,67. Según el año de residencia la concordancia interobservador osciló entre un kappa de 0,25 en los MIR de segundo año, 0,31 los de tercer año y 0,39 en los de cuarto año, con un coeficiente de Kendall de 0,79, 0,80 y 0,80 respectivamente.

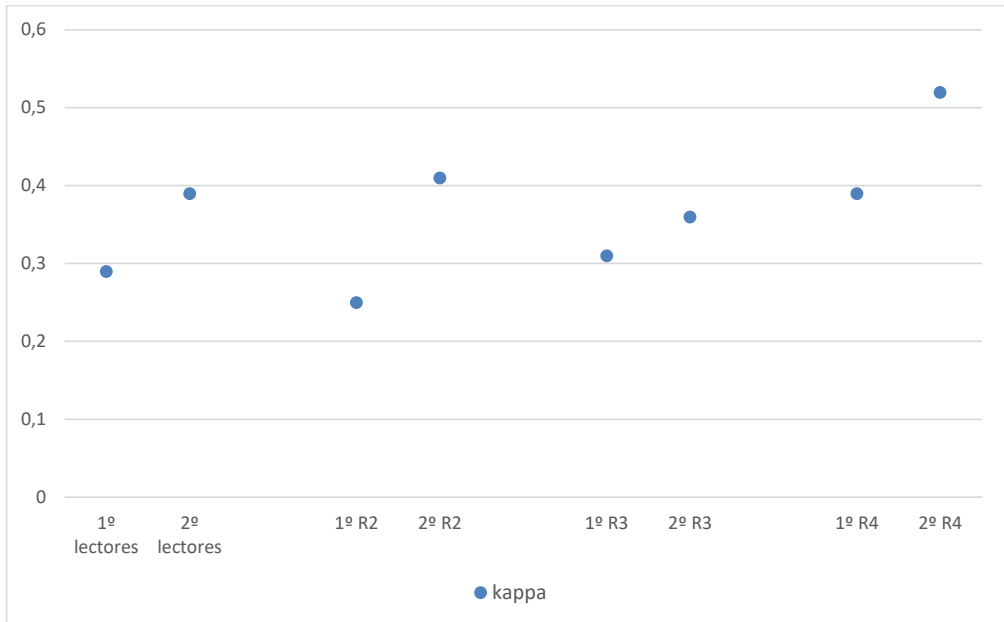
En la lectura post-formación el porcentaje de acuerdo fue de 7,27% con un IC 95% (2,00-18,00), un índice kappa global de 0,39 y un coeficiente de Kendall 0,74. Según el año de residencia, el índice kappa osciló entre 0,41 en los MIR de segundo año, 0,36 en los de tercer año y 0,52 en los de cuarto año, con un coeficiente de Kendall de 0,84, 0,80 y 0,85 respectivamente.

La diferencia entre ambas lecturas fue  $p = 0,08$ .

Los datos de variabilidad se resumen en la Tabla 20 y se representan en la Figura 22 y 23.

Tabla 20. Variabilidad interobservador pre y post-formacion determinada por el índice kappa.

	<b>Global</b>	<b>Segundo año</b>	<b>Tercer año</b>	<b>Cuarto año</b>
<b>Pre-formación</b>				
<b>% acuerdo (IC 95%)</b>	3,6(0,44-12,5)	22(11,80-35)	25,45 (15-39)	29,09(18-43)
<b>Kappa</b>	0,29±0,01	0,25±0,03	0,31±0,03	0,39±0,03
<b>Kendall</b>	0,67	0,79	0,81	0,79
<b>Post-formación</b>				
<b>% acuerdo (IC 95%)</b>	7,27 (2-17,90)	32,7(21-46,7)	32,7(21-46,7)	50,9 (37-64,5)
<b>Kappa</b>	0,39±0,01	0,41±0,035	0,36±0,035	0,52±0,035
<b>Kendall</b>	0,74	0,83	0,79	0,85
<b>p</b>	0,08	0,14	0,16	0,06

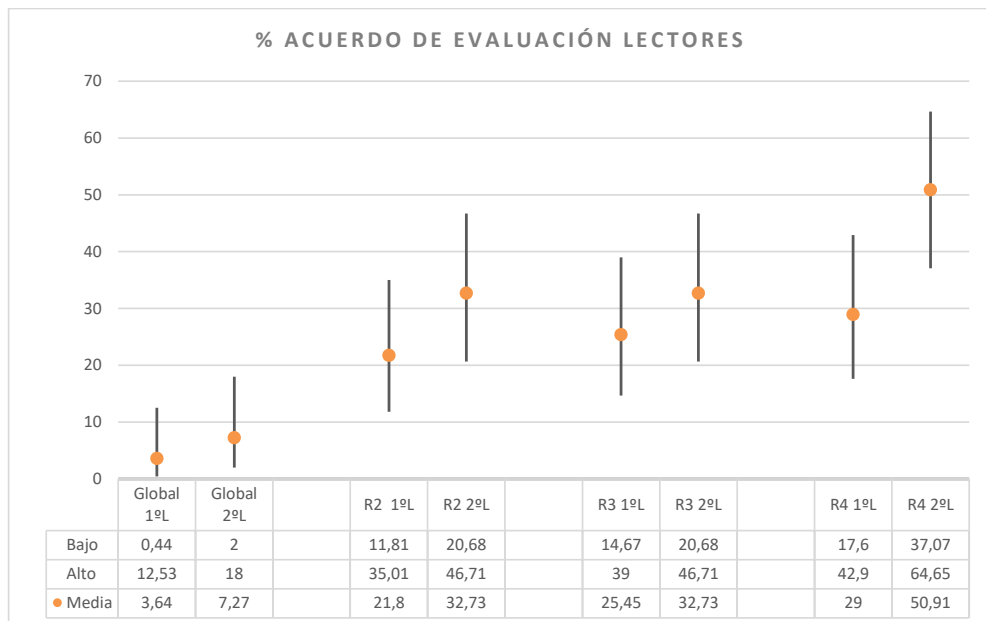


Lectores → lectores globalmente.

R2 → MIR segundo año; R3 → MIR de tercer año; R4 → MIR de cuarto año.

1º: preformación, 2º: postformación

Figura 22. Representación gráfica de los valores kappa pre y post-formación de forma global.



1ºL: lectura pre-formación; 2ºL: lectura post-formación.

R2→ MIR segundo año; R3→ MIR de tercer año; R4→ MIR de cuarto año.

Figura 23. Representación gráfica del porcentaje de acuerdos pre y postformación.

También se comparó la variabilidad interobservador de los MIR previo a la formación con el grupo de expertos de forma global. En la lectura pre-formación, los MIR obtuvieron un kappa de 0,29 con un coeficiente de Kendall de 0,67 en comparación con los lectores expertos que obtuvieron un kappa de 0,34 y 0,74 de coeficiente de Kendall ( $p = 0,08$ ).

Una vez recibida la formación específica, el grupo en formación obtuvo un porcentaje de acuerdo de 7,27% con un IC 95% (2,02-17,59) al igual que el grupo de lectores expertos, con un índice kappa de 0,39 y 0,34 respectivamente y un coeficiente de Kendall de 0,74 en ambos grupos, sin diferencias estadísticas ( $p = 0,66$ ).

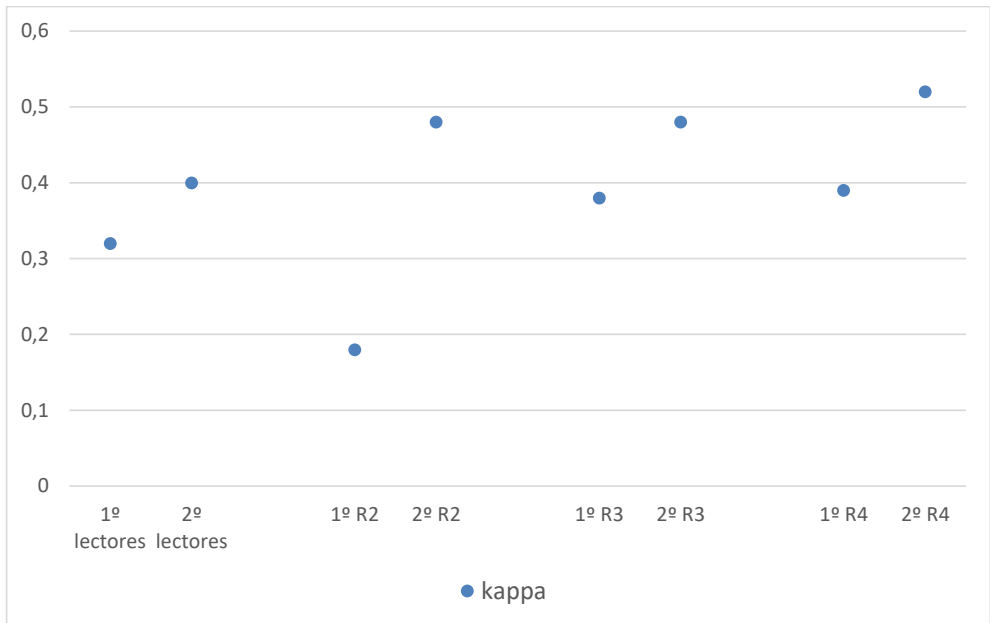
Variabilidad interobservador del grupo de lectores no expertos comparado con lector central:

Se comparó el grado de acuerdo entre los MIR en formación y el lector central. El porcentaje de acuerdo global pre-formación fue de 3,64% con un IC 95% (0,44-12,50) con un índice kappa 0,32 y un coeficiente de Kendall de 0,60. El porcentaje de acuerdo que se obtuvo post-formación fue de 7,27% con un IC 95% (2,00-18,00) con un índice kappa de 0,40 y un coeficiente de Kendall de 0,67 (p = 0,08).

Se compararon los datos en función del año de residencia, obteniendo los resultados que se reflejan en la Tabla 21 y en la Figura 24 y 25.

Tabla 21. Variabilidad interobservador de la lectura post-formación, comparada con el lector central.

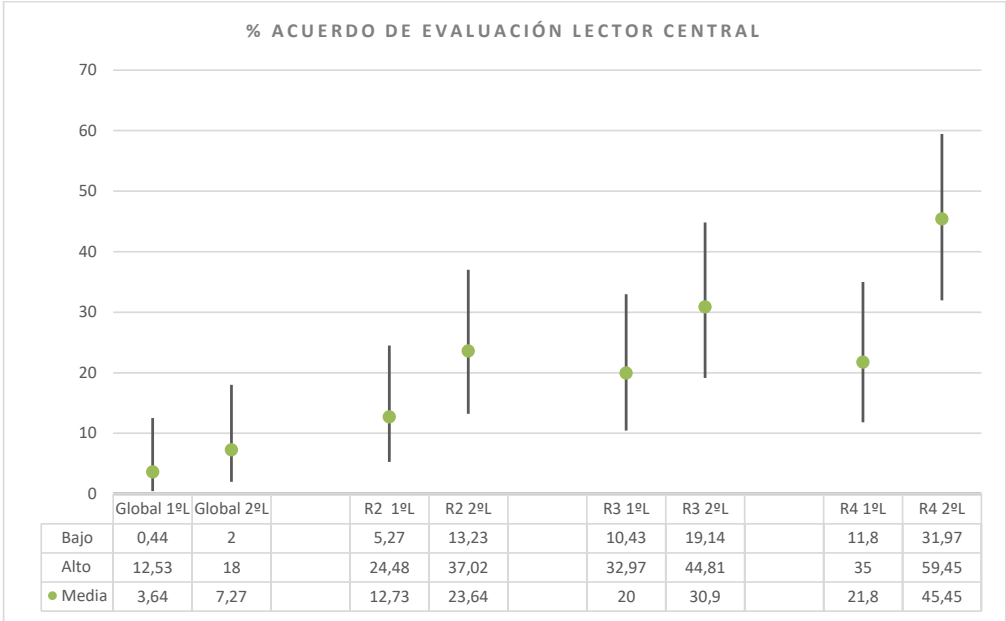
	<b>MIR Global</b>	<b>Segundo año</b>	<b>Tercer año</b>	<b>Cuarto año</b>
<b>Pre-formación</b>				
<b>% acuerdo (IC 95%)</b>	3,64(0,44-12,53)	12,70 (5-25)	20 (10,50-33)	22,00 (12-35)
<b>kappa</b>	0,32±0,02	0,18±0,04	0,38±0,04	0,39±0,04
<b>Kendall</b>	0,60	0,47	0,65	0,66
<b>Post-formación</b>				
<b>% acuerdo (IC 95%)</b>	7,27 (2-18)	23,6 (13-37)	31 (19-45)	45,4(32-59)
- <b>kappa</b>	0,49±0,02	0,48±0,04	0,48±0,04	0,53±0,04
- <b>Kendall</b>	0,67	0,66	0,64	0,71
<b>p</b>	0,08	0,139	0,21	0,07



1º: lectura pre-formación; 2º: lectura post-formación.

Lectores → lectores globalmente; R2 → MIR segundo año; R3 → MIR de tercer año; R4 → MIR de cuarto año.

Figura 24. Representación gráfica de los valores kappa pre y post-formación comparando los lectores con el lector central.



1ºL: lectura pre-formación; 2ºL: lectura post-formación.

Global → lectores globalmente; R2 → MIR segundo año; R3 → MIR de tercer año; R4 → MIR de cuarto año.

Figura 25. Representación gráfica del porcentaje de acuerdo pre y post-formación comparando los lectores con el lector central.

La variabilidad interobservador pre-formación de los MIR obtuvo un porcentaje de acuerdo 3,64%, un kappa de 0,32 con un coeficiente de Kendall de 0,59 en comparación con los lectores expertos que obtuvieron un porcentaje de acuerdo 7,27%, un kappa de 0,42 y 0,70 de coeficiente de Kendall ( $p = 0,083$ )

Tras la formación, la variabilidad interobservador de los MIR obtuvo un porcentaje de acuerdo de 7,27%, un kappa de 0,49 con un coeficiente de Kendall de 0,67, comparado con el mismo porcentaje de acuerdo, un kappa 0,42 y el Kendall de 0,7, obtenido por los expertos ( $p = 0,71$ ).

Variabilidad del índice de Rutgeerts modificado agrupado.

La variabilidad del índice de Rutgeerts modificado agrupado queda reflejada en la Tabla 22. Y la variabilidad del índice de Rutgeerts modificado agrupado del grupo de lectores no expertos comparado con el lector central, se resume en la Tabla 23.



Tabla 22. Variabilidad interobservador del índice de Rutgeerts modificado agrupado, pre y post-formación.

	MIR global	R2	R3	R4
<b><math>i &lt; 2</math> vs <math>\geq 2</math></b>				
<b>Pre-formación</b>				
%acuerdo	49,09	74,55	72,73	69,09
IC 95%	(35,35-62,93)	(61,00-85,33)	(59,04-83,86)	(55,19-80,86)
Kappa $\pm$ DE	0,54 $\pm$ 0,02	0,58 $\pm$ 0,08	0,55 $\pm$ 0,08	0,55 $\pm$ 0,08
<b>Postformación</b>				
%acuerdo	54,55	80,00	63,64	82,27
IC 95%	(40,55-68,03)	(67,03-89,57)	(49,56-76,19)	(75,52-94,73)
Kappa $\pm$ DE	0,62 $\pm$ 0,02	0,68 $\pm$ 0,08	0,38 $\pm$ 0,08	0,77 $\pm$ 0,08
<b>p</b>	0,19	0,70	0,70	0,20
<b><math>i \leq 2a</math> vs <math>\geq 2b</math></b>				
<b>Preformación</b>				
%acuerdo	32,73	58,18	63,64	58,18
IC 95%	(20,68-46,71)	(44,11-71,35)	(49,56-76,19)	(44,11-71,35)
Kappa $\pm$ DE	0,44 $\pm$ 0,02	0,44 $\pm$ 0,08	0,51 $\pm$ 0,08	0,44 $\pm$ 0,08
<b>Postformación</b>				
%acuerdo	45,45	58,18	70,91	81,82
IC 95%	(31,97-59,45)	(44,11-71,35)	(57,10-82,37)	(69,10-90,92)
Kappa $\pm$ DE	0,54 $\pm$ 0,02	0,44 $\pm$ 0,08	0,60 $\pm$ 0,08	0,72 $\pm$ 0,08
<b>p</b>	0,02	1,00	0,70	0,02
<b><math>i \leq 2</math> vs <math>\geq 3</math></b>				
<b>Pre-formación</b>				
%acuerdo	27,10	56,36	56,36	61,82
IC 95%	(16,14-40,96)	(42,32-69,70)	(42,32-69,70)	(47,73-74,59)
Kappa $\pm$ DE	0,34 $\pm$ 0,02	0,32 $\pm$ 0,08	0,37 $\pm$ 0,08	0,38 $\pm$ 0,08
<b>Postformación</b>				
%acuerdo	43,64	72,73	72,73	69,09
IC 95%	(30,3-57,70)	(59,04-83,86)	(59,04-83,86)	(55,19-80,86)
Kappa $\pm$ DE	0,50 $\pm$ 0,02	0,45 $\pm$ 0,08	0,06 $\pm$ 0,08	0,58 $\pm$ 0,08
<b>p</b>	0,007	0,20	0,20	0,70

Tabla 23. Variabilidad interobservador del índice de Rutgeerts modificado agrupado pre y post formación comparado con el lector central.

	MIR	R2	R3	R4
<b><math>i \leq 1</math> vs <math>\geq 2</math></b>				
<b>Preformación</b>				
%acuerdo	47,27	69,09	70,91	65,45
IC 95%	(33,65-61,03)	(55,19-80,86)	(57,10-82,37)	(51,42-77,76)
Kappa±DE	0,60±0,05	0,55±0,08	0,65±0,08	0,59±0,08
<b>Postformación</b>				
%acuerdo	54,55	78,18	63,64	82,27
IC 95%	(40,55-68,03)	(64,99-88,19)	(49,56-76,19)	(75,52-94,73)
Kappa±DE	0,74±0,05	0,77±0,08	0,62±0,08	0,84±0,08
<b>p</b>	0,26	0,70	0,70	0,20
<b><math>i \leq 2</math> vs <math>\geq 2b</math></b>				
<b>Preformación</b>				
%acuerdo	32,73	43,64	54,55	50,91
IC 95%	(20,68-46,71)	(30,30-57,68)	(40,55-68,03)	(37,07-64,65)
Kappa±DE	0,41±0,04	0,31±0,07	0,48±0,08	0,44±0,08
<b>Postformación</b>				
%acuerdo	45,45	50,91	62,27	76,36
IC 95%	(31,97-59,45)	(37,07-64,65)	(53,29-79,32)	(62,98-86,77)
Kappa±DE	0,61±0,04	0,50±0,08	0,63±0,08	0,67±0,08
<b>p</b>	0,02	0,70	0,70	0,70
<b><math>i \leq 2</math> vs <math>\geq 3</math></b>				
<b>Preformación</b>				
%acuerdo	27,27	47,27	54,55	61,82
IC 95%	(16,14-40,96)	(33,65-61,20)	(40,55-68,03)	(47,73-74,59)
Kappa±DE	0,47±0,04	0,33±0,08	0,55±0,08	0,55±0,08
<b>Postformación</b>				
%acuerdo	43,64	63,64	67,27	65,45
IC 95%	(30,30-57,68)	(49,56-76,19)	(53,29-79,32)	(51,42-77,76)
Kappa±DE	0,56±0,02	0,47±0,08	0,60±0,08	0,58±0,08
<b>p</b>	0,007	0,20	0,20	0,70

\*DE: Desviación estándar

Otros análisis incluidos.

También se comparó la variabilidad pre y postformación de los lectores no expertos con los expertos. Tablas 24 y 25.

Tabla 24. Diferencias entre la lectura pre-formación de lectores no expertos comparada con el grupo de expertos globalmente en las distintas agrupaciones.

	<b>Expertos</b>	<b>No expertos</b>	<b>p</b>
<b>i&lt;2 vs i≥2</b>			
% acuerdo IC 95%	50,91 (37,07-64,65)	49,09 (35,35-62,93)	0,41
kappa±DE*	0,56±0,03	0,54±0,02	
<b>i≤2a vs i≥2b</b>			
% acuerdo IC 95%	36,36 (23,81-50,44)	32,73(20,68-46,71)	0,33
kappa±DE*	0,49±0,03	0,44±0,02	
<b>i≤2 vs i≥3</b>			
% acuerdo IC 95%	43,64 (30,30-57,68)	27,27 (16,14-40,96)	0,0069
kappa±DE*	0,43±0,03	0,34±0,02	

\*DE: Desviación estándar

Tabla 25. Diferencias entre la lectura post-formación de lectores no expertos comparada con los expertos.

	<b>Expertos</b>	<b>No expertos</b>	<b>p</b>
<b><math>i &lt; 2</math> vs <math>i \geq 2</math></b>			
% acuerdo IC 95%	50,91 (37,07-64,65)	54,55 (40,55-68,03)	0,36
Kappa $\pm$ DE*	0,56 $\pm$ 0,03	0,62 $\pm$ 0,02	
<b><math>i \leq 2a</math> vs <math>i \geq 2b</math></b>			
% acuerdo IC 95%	36,36 (23,81-50,44)	45,45 (31,97-59,45)	0,16
kappa $\pm$ DE*	0,49 $\pm$ 0,03	0,54 $\pm$ 0,02	
<b><math>\leq 2</math> vs <math>i \geq 3</math></b>			
% acuerdo IC 95%	43,64 (30,30-57,68)	43,64 (30,30-57,68)	0,96
kappa $\pm$ DE*	0,43 $\pm$ 0,03	0,49 $\pm$ 0,02	

\*DE: Desviación estándar

### 8. Vídeos con puntuaciones diferentes.

Los vídeos que más difirieron en las puntuaciones otorgadas, en comparación con el lector central, fueron los catalogados como i4 e i2b. En la Figura 26 queda reflejado el número de vídeos de cada categoría que en comparación con el lector central, se puntuaron de diferente forma. Se puede observar cómo en la categoría i2b e i4 son las que registran más puntuaciones diferentes.

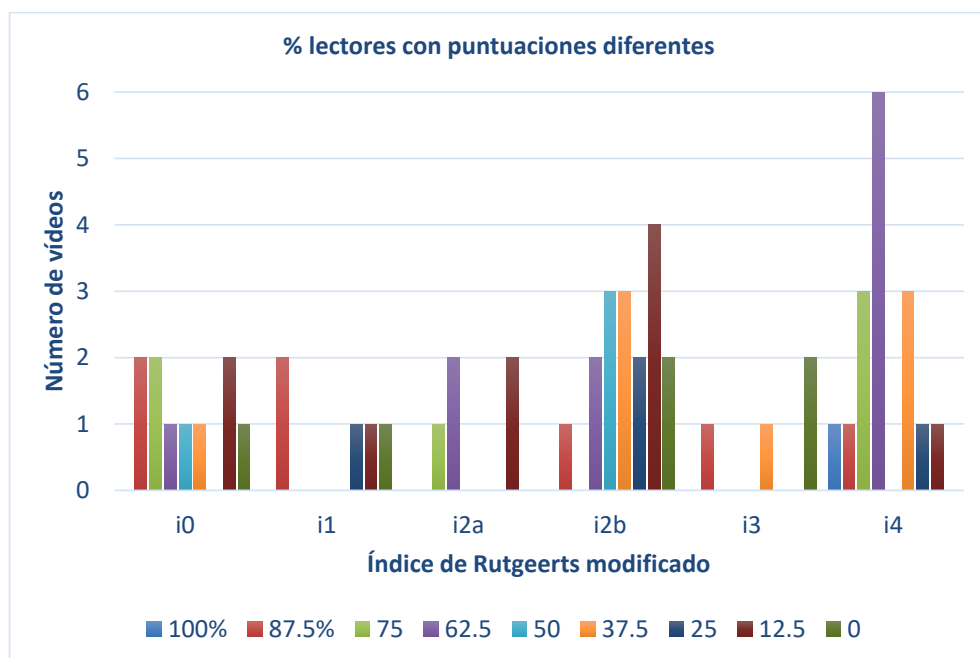


Figura 26. Porcentaje de lectores que difirieron en las puntuaciones de cada categoría del índice de Rutgeerts en comparación con el lector central.



## 9. DISCUSIÓN

---





La endoscopia tiene un interés creciente y un papel cada vez más relevante en las enfermedades digestivas, incluyendo la EII. Hoy en día resulta imprescindible disponer de una valoración lo más objetiva posible sobre la actividad inflamatoria para poder instaurar los tratamientos más adecuados y realizar estrategias adaptadas a cada paciente.

Desde hace años se conoce que no existe una buena correlación entre la clínica de los pacientes y la actividad inflamatoria. Esto, añadido al interés creciente por la curación mucosa como objetivo terapéutico, no sólo para mejorar clínicamente al paciente sino para intentar modificar el curso evolutivo de la enfermedad, ha supuesto que la endoscopia sea un pilar fundamental en la EII y sobre todo en la EC.

No obstante, la endoscopia para valorar la curación mucosa en la EC es limitada debido a dos aspectos fundamentales, uno de ellos por su afectación transmural y otro, por la posibilidad de afectar cualquier tramo intestinal y no ser accesible mediante técnicas endoscópicas. A pesar de ello, debido a que la afectación ileocólica es la más frecuente y la recurrencia de la enfermedad en los pacientes intervenidos ocurre en el neofleon, la colonoscopia sigue siendo el patrón oro para valorar la actividad.

En la valoración objetiva de la actividad inflamatoria hay dos aspectos fundamentales, por una parte, la **herramienta** disponible y por otra la **persona** encargada de realizar esa valoración. Ambos aspectos van a ser fundamentales para conseguir una evaluación lo más objetiva posible.

### **Herramienta: índice de Rutgeerts.**

Las herramientas disponibles son los índices endoscópicos, de los cuales sólo uno de ellos se centra en valorar la RPQ endoscópica, el índice de Rutgeerts<sup>100,101</sup> o su modificación propuesta en 2008 por Domenech et al<sup>145</sup>.

En 1984 cuando la recurrencia de la enfermedad era un aspecto prácticamente desconocido, Rutgeerts et al<sup>101</sup> publicó un artículo sobre la historia natural de la misma. En ese artículo estableció una definición de úlcera aftosa y úlcera profunda, aspectos fundamentales en la

clasificación de este índice, considerando RPQ endoscópica la aparición de lesiones neoleales en tramos previamente sanos incluso con biopsia<sup>100</sup>. Posteriormente, en 1990, demostró que los pacientes con lesiones más graves presentaban un curso más agresivo de la enfermedad con necesidad de reintervenciones en muchos de ellos<sup>101</sup>. Sin embargo, este índice no está validado y además se definió en una muestra muy pequeña de pacientes con el objetivo de valorar el seguimiento inmediato postoperatorio. Por tanto, no se creó como una herramienta para realizar un seguimiento endoscópico rutinario.

A pesar de todo ello, esta clasificación se utilizó de forma universal y no se cuestionó hasta que en 2008 Domenech et al<sup>145</sup> realizaron una subclasificación de una de las categorías del índice de Rutgeerts, i2, ya que observaron que la evolución en este grupo era heterogénea y esto podría deberse a que no todas las lesiones clasificadas en este subgrupo se debían a la actividad inflamatoria propiamente dicha sino a fenómenos isquémicos<sup>145</sup>. A pesar de que esta subclasificación no se ha utilizado tan extensamente, en los últimos años también se ha cuestionado. Dos estudios retrospectivos han demostrado cómo la evolución entre i2a e i2b no difiere en términos de RPQ clínica ni en la necesidad de tratamiento endoscópico o quirúrgico, incluso cambiando la medicación inmunomoduladora<sup>146,147</sup>

Tanto el índice de Rutgeerts original como el modificado, a pesar de ser índices sencillos de utilizar, incluyen aspectos que pueden resultar confusos y ser los responsables de la variabilidad en la puntuación de la actividad inflamatoria.

En primer lugar, el tramo intestinal definido como neofleon. Inicialmente la recurrencia se definió como la afectación exclusiva de, aproximadamente, los últimos 15cm de neofleon, lo cual en ocasiones resulta complicado de medir con exactitud. El tramo explorado es importante ya que una afectación más proximal no debería considerarse recurrencia sino afectación proximal propiamente dicha, ocasionando un cambio en la clasificación fenotípica de la enfermedad, L4, sin incluir sus hallazgos en el índice de Rutgeerts o en su versión modificada.

En segundo lugar, la definición de algunas variables incluidas en el índice, como son la diferencia entre una úlcera superficial y profunda, la afectación de la mucosa subyacente

(fundamentalmente el edema de la mucosa) o la presencia de lesiones salteadas entre mucosa normal.

La mayoría de los estudios de variabilidad y concordancia se han realizado con los índices endoscópicos de actividad luminal. Si se extrapolan los datos obtenidos en los estudios realizados con el CDEIS, se puede observar cómo el edema de la mucosa no es una variable incluida en este índice, por estar íntimamente asociada con las ulceraciones, y cómo la profundidad de la úlcera fue la variable responsable del mayor grado de discordancia entre los lectores. Por este motivo al elaborar el SES-CD, se sustituyó la profundidad por el tamaño de la úlcera, que es un parámetro más fácil de calcular durante la endoscopia, para evitar o minimizar esa variabilidad. De momento, sólo se ha publicado un estudio en el que se analiza la variabilidad interobservador de la profundidad de las úlceras obteniendo una concordancia global con un kappa de 0,12 vs 0,40 según si se clasifican como úlceras aftosas o úlceras profundas, confirmando los datos obtenidos previamente con el CDEIS<sup>220</sup>.

En tercer lugar, en las categorías del índice de Rutgeerts y en su versión modificada se incluyó un concepto que es la presencia de lesiones grandes salteadas sobre áreas de mucosa normal, que se desconoce si son úlceras, pseudopólipos o qué tipo de lesiones, ya que no está descrito ni recogido en esos artículos iniciales.

Por último, el tipo de anastomosis. Inicialmente, cuando Rutgeerts describió el índice, todos los pacientes intervenidos con una resección ileocólica o ileocecal presentaban anastomosis termino-terminales. Sin embargo, con la evolución de las técnicas quirúrgicas, actualmente no sólo se encuentran las anastomosis terminotermiales sino las anastomosis laterolaterales, de las cuales hay dos tipos (isoperistálticas y anisoperistálticas). En estas anastomosis siempre queda un fondo ciego ileal y un pliegue interasas de mucosa intestinal. Estrictamente no se considera como neóleon y la afectación de esta zona puede estar relacionada con el sobrecrecimiento bacteriano o bien con fenómenos isquémicos. Sin embargo, son hipótesis, ya que en la descripción inicial no se planteó este problema puesto que las anastomosis eran todas termino-terminales.

Ante la dificultad en interpretar correctamente el tramo de neóleon y las lesiones visualizadas, es de esperar que la variabilidad del índice de Rutgeerts original o modificado

entre los observadores sea elevada. Debido a ello, a pesar de ser el único índice disponible se consideraría una herramienta poco fiable para determinar la gravedad y totalmente supeditada a la persona que realiza la lectura de la endoscopia. No obstante, la información disponible al respecto es escasa y en ocasiones contradictoria.

### **Lectores:**

Debido a lo referido previamente, para minimizar esa variabilidad entre los lectores, en los últimos años ha aparecido una nueva figura, el lector central o el *offsite reader*<sup>221</sup>, como una posible solución a este problema. A pesar de la ventaja que ha supuesto en otros ámbitos, todavía no se dispone de suficientes datos como para utilizar esta figura de forma universal en la endoscopia digestiva. *A priori*, se trata de una persona experta, ciega a la información sobre el paciente y con experiencia no sólo en la EII sino también en el uso de los índices endoscópicos, lo que implica un criterio uniforme de puntuación. Sin embargo, no existen unos criterios preestablecidos que deba cumplir de una forma estricta o de un curso de formación que deba superar para convertirse en lector central. De esta forma, la elección suele ser basada en la experiencia.

Aunque hay razones teóricas por las que se espera que su lectura mejore la exactitud de la valoración endoscópica, por sí solo no reduce la variabilidad en la lectura<sup>222</sup>. Hay otros factores que también influyen, como son la calidad de las imágenes y los sistemas de votación o los esquemas de lectura, en los que se pueden combinar los lectores centrales con los lectores locales.

La inclusión de los lectores locales en el panel de lectura puede suponer una ventaja, ya que inevitablemente se encuentran involucrados en los estudios por ser los encargados de grabar los vídeos y procesarlos en la mayor calidad posible para poder permitir la lectura correcta por parte de un lector independiente. Esto es una diferencia fundamental con la radiología, en estos casos todos los lectores valoran las mismas imágenes mientras que en el ámbito de la endoscopia, los lectores centrales van a valorar las imágenes o los vídeos que escogen los lectores locales.

Por ello, en referencia a la lectura, se pueden plantear diferentes esquemas, exclusivamente con lectores centrales o bien, esquemas mixtos con lectores locales y centrales o exclusivamente con lectores locales. En estos esquemas mixtos el lector local realiza la lectura conforme se produce el procedimiento, y posteriormente realizaría la lectura un lector central independiente. Al incluir lectores locales o no expertos surge la necesidad de formar adecuadamente a estos lectores en la utilización de los índices endoscópicos para valorar la actividad inflamatoria.

El problema aparece cuando, en los esquemas mixtos, se observan discrepancias entre ambos. En estos casos se pueden resolver de dos formas, o bien por consenso entre los lectores o bien incluyendo a más lectores centrales, ciegos a las lecturas previas. Con este último sistema, llamado *Voting for Image Scoring and Assessment (VISA)*, habitualmente se proponen dos lectores siendo uno de ellos el lector local y el otro el lector central. En caso de discrepancia entre ambos, se incluye a un tercer lector, que es otro lector central, coincidiendo en la mayoría de los casos. Este sistema supone un aumento de coste y de tiempo del estudio<sup>222</sup>.

Se han establecido varios esquemas de lectura, con sus métodos para resolver discrepancias<sup>221,223</sup>. A modo de resumen se pueden establecer siete posibilidades (Tabla 26).

Tabla 26. Esquemas de lectura ( *Adaptada de Panes et al, JCC 2016*)<sup>221</sup>.

Esquema de lectura	Lector central	Lector local	Método para resolver discrepancias
1	1	Ninguno	--
2	1 excepto vídeos de mala calidad, se incluye otro.	Ninguno	Desacuerdo Lector central 1 y lector central 2= consenso.
3	Hasta 2	1	Desacuerdo lector central 1 y lector local → lector central 2. Acuerdo lectores centrales → puntuación final. Desacuerdo lectores centrales → consenso.
4	2	Ninguno	Igual que el esquema 2
5	Hasta 3	1	Acuerdo lector local y lector central → puntuación final. Desacuerdo Lector central 1 y lector local → valoración lector central 2. Acuerdo lectores centrales → puntuación final. Desacuerdo Lector central 1/2 → valoración lector central 3. Puntuación final: consenso entre lectores centrales.
6	Hasta 3	Ninguno	Acuerdo lectores centrales 1/2 → puntuación final. Desacuerdo lectores centrales 1/2 → lector central 3. Puntuación final: mayoritaria entre los tres lectores centrales

Esquema de lectura	Lector central	Lector local	Método para resolver discrepancias
7	2+lector central adjudicador	ninguno	Acuerdo lectores centrales 1 y 2 → puntuación final Desacuerdo lectores centrales 1 y 2 → se utiliza la puntuación del lector central 2

De momento, en la endoscopia digestiva no existen guías de práctica clínica en las que se base la FDA o la EMA para solventar estas dudas sobre los esquemas de lectura<sup>185,224</sup>.

Por una parte, la FDA propone que la centralización de las imágenes no es crucial y va a depender de varios aspectos: que se asegure que el lector es ciego a la información del paciente, que se prevea una variabilidad en la lectura de las imágenes y que los sesgos derivados de la interpretación de las mismas estén controlados con el diseño del estudio<sup>184</sup>. Aunque también refiere que los procesos centralizados aseguran la formación continua del lector, lo que ayuda a garantizar la calidad de las imágenes y reducir la variabilidad consiguiendo una estimación más precisa del efecto.

Por otra parte, la EMA recomienda que deben ser dos o más *off-site reader* los que evalúen las imágenes, pero de forma independiente, no en consenso<sup>224</sup>.

La EMA advierte que los paneles de consenso no deben utilizarse como primer método de evaluación ya que siempre hay un miembro más influyente que otro. Los lectores deben ser independientes, no verse influenciados por otros lectores y ser ciegos al estudio. Además, no deberían utilizarse sólo imágenes por la clara desventaja con respecto a los lectores locales que disponen de video.

No obstante, esto se basa en opiniones de expertos ya que no hay comparación entre distintos esquemas de lectura que muestren la superioridad de uno frente a otro.

Por tanto, el *off-site reader* es una figura experta que puede aportar una mayor objetividad a la lectura de los vídeos. Sin embargo, todavía se desconoce qué profesional debe ser, qué

criterios que debe cumplir, cuál es el número más adecuado de lectores o cómo se debería componer el esquema de lectura para que sea lo más eficiente posible.

En resumen, por un lado disponemos de un único índice, que por las variables que incluye y sus características intrínsecas, tiende a inducir variabilidad en su interpretación. Y por otro lado, existen varias alternativas para poder minimizar esta variabilidad, que se centran en el sistema de lectura y concretamente en la combinación de lectores. No obstante, la evidencia al respecto es escasa y los estudios publicados muestran resultados discordantes. Debido a esto, se propuso este estudio, cuyo objetivo fue conocer la variabilidad interobservador de dos grupos de lectores expertos como son los expertos en endoscopias y los expertos clínicos en EII, y si la inclusión de un lector central podría ofrecer beneficios en términos de concordancia. Además, también nos propusimos valorar si la formación de lectores no expertos como podrían ser lectores locales, podría mejorar la variabilidad interobservador en este grupo equiparándolo al grupo de expertos, siendo un punto a tener en cuenta en el futuro de la formación en gastroenterología.

En la Figura 15, página 117 se resumen las comparaciones realizadas durante el estudio.

### 9.1. Variabilidad intraobservador.

En el estudio que presentamos, se ha determinado la variabilidad intraobservador en un grupo de lectores expertos en diferentes campos, en la endoscopia y en la EII. Este valor es crucial para determinar la repetitividad del índice. Teóricamente, si el índice presenta una definición clara y concisa, los lectores deben realizar la lectura siempre de la misma forma. Sin embargo, los resultados obtenidos ofrecen dudas al respecto.

Los estudios publicados hasta la fecha se basan mayoritariamente en la variabilidad interobservador. Los únicos dos estudios en los que se valora la variabilidad intraobservador de lectores centrales obtuvieron una elevada concordancia con un CCI de 0,83 (IC 95% 0,75-0,90) y 0,85 (IC 95% 0,78-0,89) respectivamente<sup>179,220</sup>.

En este apartado destacan dos puntos. En primer lugar, en nuestro estudio el lector central fue el lector que peor concordancia intraobservador obtuvo con un kappa de 0,53 y un



coeficiente de Kendall de 0,64. Con estos datos se planteó la duda de si realmente fue el lector más adecuado para considerarlo como lector central, a pesar de cumplir todos los criterios descritos por la literatura. El único dato a favor del lector central escogido fue que, al comparar el porcentaje de acuerdo, los lectores expertos mostraron un mayor acuerdo al compararlo con este lector que con los otros dos lectores centrales propuestos (experto en EII y otro lector escogido aleatoriamente), de forma estadísticamente significativa ( $p = 0,012$ ). No obstante, la duda queda patente y pone de manifiesto que probablemente la experiencia clínica y la participación en ensayos clínicos no sean los mejores criterios para elegir a un lector central, sino que se debería basar en datos más objetivos. Una solución podría ser realizar de forma periódica análisis de variabilidad intraobservador en los lectores considerados expertos y asegurar su actualización en el uso de índices endoscópicos. De esta forma, en el momento que se plantea un estudio de variabilidad interobservador, se disponen de datos objetivos para escoger a un lector central que muestra de forma periódica una buena variabilidad intraobservador.

En segundo lugar, se obtuvieron resultados discordantes según el grupo de expertos. Los expertos en endoscopias fueron los que mejores datos de concordancia obtuvieron con un kappa de 1,00 a diferencia de los expertos clínicos en EII, donde hubo más variabilidad (kappa entre 0,56 -1,00). A pesar de la desigualdad en el número de lectores expertos en cada grupo, la diferencia en el grado de acuerdo que obtuvieron los dos grupos de expertos podría explicarse debido a la costumbre que tienen los expertos en endoscopias de utilizar índices o escalas que incluyen la interpretación de imágenes endoscópicas. Los expertos en endoscopias están más habituados a puntuar las lesiones endoscópicas utilizando siempre los mismos criterios, independientemente de otros factores inherentes al paciente o a la exploración. Por el contrario, en el caso de los expertos en EII, durante la evaluación de los pacientes en consultas no se rigen por una única escala, sino que la toma de decisiones varía en función de multitud de datos (los datos endoscópicos, los radiológicos, los analíticos y los clínicos). Aunque en este caso todos los lectores fueron ciegos a la información clínica del paciente, por su actividad diaria es más probable que los expertos en endoscopias clasifiquen las lesiones siempre de la misma forma.

Al analizar el grupo de clínicos expertos en EII se observaron datos discrepantes e incluso contradictorios entre sus lectores. El lector que obtuvo una concordancia más baja ( $\kappa = 0,56$ ) clasificó de la misma forma los extremos de la enfermedad (i1-i4) discrepando fundamentalmente en la subcategoría i2. Sin embargo, el coeficiente de Kendall fue superior con cifras de 0,86 lo que implicó que, aunque las puntuaciones centrales no fueran exactamente las mismas, se encontraban muy próximas entre ellas. El otro lector presentó una concordancia superior determinada por un índice  $\kappa$  de 0,75, puntuando de la misma forma las categorías intermedias y discrepando en los extremos. Sin embargo, en este caso el coeficiente de Kendall fue inferior, 0,69, lo que implicó que en aquellos vídeos que no puntuó exactamente de la misma forma, se desvió más de la puntuación otorgada inicialmente. Realmente con estos resultados no podemos extraer unas conclusiones concretas ya que sólo fueron dos lectores los que obtuvieron menor variabilidad intraobservador y con datos discordantes, pero refleja cómo el índice  $\kappa$ , no es el mejor indicador para valorar esta variabilidad.

A pesar de lo expuesto previamente, todos los lectores expertos obtuvieron un coeficiente de Kendall adecuado (superior a 0,70), sin embargo, sólo en los lectores expertos en endoscopias se pudo considerar el índice de Rutgeerts modificado como un índice repetitivo.

## 9.2. Variabilidad interobservador.

Otro de los factores fundamentales para un índice es la reproducibilidad, es decir, que independientemente de quien realice la lectura, siempre se debería puntuar de la misma forma las lesiones, lo que le confiere robustez al índice. Esta es una de las características que más interesa debido precisamente a que no siempre es la misma persona la que realiza la endoscopia y valora la gravedad de las lesiones visualizadas. En este trabajo se han incluido a un grupo de expertos, y a un grupo de no expertos, que son los MIR.

Actualmente no existe una definición clara y bien establecida entre experto-no experto. En algunos estudios se considera la experiencia en la endoscopia de la EII<sup>181</sup>, en otros la experiencia en utilizar los índices endoscópicos en la EII y la participación en ensayos

clínicos<sup>175</sup> y en otros casos deben cumplir varios criterios ( consulta monográfica viendo más de 40 pacientes de EII/mes; realizar más de 20 exploraciones mensuales a pacientes con EII y utilizar los índices endoscópicos en todos los informes de endoscopias realizadas)<sup>182</sup>. No obstante, no suelen valorar la diferencia entre varios grupos de expertos. Los lectores que se incluyeron como expertos cumplían los criterios utilizados en los estudios publicados hasta ahora, tanto clínicos como endoscópicos por separado, por ello los resultados obtenidos fueron comparables.

Antes de comparar los resultados con la bibliografía disponible, en el trabajo que presentamos destaca un aspecto fundamental que es el que va a justificar los datos de concordancia obtenidos tanto en el índice de Rutgeerts modificado como en el agrupado. Según se puede observar en la figura 16 de la página 127 la mayoría de vídeos puntuados por los lectores expertos pertenecen a la categoría i2 (40%) y también hay muchos vídeos puntuados como i4 (17%). Como hemos comentado previamente el índice de Rutgeerts incluye dos variables que resultan complicadas de diferenciar como son el edema de mucosa, es decir la identificación de la ileitis, y la profundidad de la úlcera (diferenciar úlceras aftosas de úlceras profundas). Pues bien, precisamente estos dos conceptos son los que van a definir y diferenciar a las categorías i2b e i4. Por ello los datos de concordancia de este trabajo, como veremos a continuación, son inferiores a los publicados hasta ahora. De hecho, en la figura 26 de la página 149 queda reflejado cómo los vídeos que más discrepancia presentan en su puntuación son precisamente los que pertenecen a la categoría i2b e i4.

La concordancia del índice de Rutgeerts modificado obtenida en nuestro trabajo en el grupo global de lectores expertos es pobre ( $\kappa = 0,34$ ), inferior a la descrita en los estudios publicados utilizando el índice de Rutgeerts tanto original como modificado. Hasta ahora la mayoría de información de la que disponemos se centra en el índice de Rutgeerts original, en el que la concordancia entre expertos o lectores centrales es moderada con un  $\kappa$  que varía entre 0,42 y 0,57<sup>175,180-182</sup>.

El motivo de este rango en el grado de concordancia en los diferentes trabajos publicados, incluido el nuestro, no radica en el número de visualizaciones sino más bien en el número de colonoscopias que están incluidas en cada categoría. En el estudio de Marteau et al<sup>181</sup>,

donde la variabilidad interobservador es 0,43, prácticamente la mitad de las colonoscopias incluidas tenían una valoración i2 y sólo una minoría (10%) fueron catalogadas como i3-i4. En nuestro trabajo ocurre lo mismo, la puntuación mayoritaria otorgada a los vídeos es i2 (40%), incluyendo i2a e i2b, con porcentajes que varían de 20-67% según el lector (figura 16, página 127). En otros estudios en los que obtienen mayor concordancia como es el de Kennedy et al<sup>180</sup> con un kappa de 0,50, ocurre al contrario, la mayoría de colonoscopias son i3-i4 (35% de los casos) y sólo el 16% son i2.

En los únicos dos estudios en los que utilizan el índice de Rutgeerts modificado, obtienen mejores datos de concordancia que en nuestro trabajo e incluso que en los trabajos en los que se utiliza el índice de Rutgeerts original. En uno de ellos, Gecse et al<sup>179</sup> incluyeron a cuatro lectores centrales que valoraron 30 exploraciones sin especificar la puntuación de cada colonoscopia, mostrando datos similares a los publicados por Kennedy et al<sup>180</sup>. Aunque no podemos saber el motivo de esa buena concordancia, probablemente se deba a las categorías de los vídeos incluidos. En otro estudio, más reciente<sup>220</sup>, se incluyeron a 4 lectores centrales y obtuvieron un CCI de 0,71 en el índice de Rutgeerts original y 0,74 en el modificado. En este caso, la buena concordancia también se relacionó con los vídeos incluidos en la categoría i2 ( un 30%, siendo mayoritariamente i2a). En nuestro trabajo, a pesar de que la puntuación mayoritaria fue i2, prácticamente la mitad era i2a y la otra mitad i2b. Esto justifica que realmente las lesiones localizadas en la anastomosis no son las responsables de esa variabilidad sino las localizadas en el neóleon (i2b) y por ello en nuestro trabajo se obtuvo una concordancia menor.

En ninguno de los trabajos mencionados se calculó el coeficiente de Kendall por lo que no es posible comparar los resultados ni establecer conclusiones al respecto. Si bien es cierto, los lectores tanto de forma global como por grupo de expertos obtuvieron una buena concordancia con cifras superiores a 0,75.

Por tanto, parece claro que la categoría i2 y posiblemente la subcategoría i2b sea la que genera la mayoría de las discrepancias en la puntuación del índice de Rutgeerts modificado justificando esa concordancia pobre del índice (kappa = 0,34), inferior a otros trabajos. Por ello, y ante la dificultad de clasificar las lesiones en una categoría concreta, también se planteó la posibilidad de agrupar varias categorías en función de diferentes definiciones de

recurrencia. No es tan importante incluir al paciente en una categoría concreta o en otra sino conocer si presenta o no recurrencia para poder establecer una estrategia terapéutica adecuada.

En nuestro trabajo se establecieron tres agrupaciones en función de varios conceptos, dos definiciones de RPQ endoscópica ( $\geq i2$  y  $\geq i2b$ ) y una definición de RPQ endoscópica grave ( $\geq i3$ ). Independientemente de la definición que se establezca, parece que hay unanimidad en que la concordancia del índice agrupado mejora de forma sustancial, en comparación con el índice no agrupado.

En la mayoría de los estudios publicados optan por el punto de corte en  $\geq i2$  para considerar RPQ endoscópica. De esta forma, la concordancia entre expertos del grupo  $< i2$  vs  $\geq 2$ , varía de CCI de 0,44 (IC 95% 0,24-0,65) en el grupo de Fernandes<sup>182</sup>, kappa de 0,47 en el estudio de Marteau<sup>181</sup> o kappa de 0,83 en el de Kennedy<sup>180</sup>. En nuestro trabajo obtenemos unos datos muy similares con un kappa de 0,56, debido a que precisamente el grupo mayoritario en puntuaciones es el  $i2$ , seguido del  $i4$ , que se encuentran fusionados en esta agrupación.

Al utilizar el índice de Rutgeerts modificado, planteamos otra opción de definición de RPQ endoscópica como son las puntuaciones  $\geq 2b$ , asumiendo que  $i2a$  son lesiones isquémicas y por tanto no están relacionadas con la enfermedad. En este caso la concordancia fue discretamente inferior a la obtenida en la primera agrupación con un kappa de 0,49, lo que se puede justificar por la propia subdivisión de la categoría  $i2$ , confirmando que esta categoría es el origen de la discrepancia y siendo el subgrupo  $i2b$  el responsable de la misma. Además, concretamente en esta agrupación sí que se observaron diferencias muy próximas a la significación estadística entre los dos grupos de expertos a favor de los endoscopistas (kappa = 0,61 vs kappa = 0,45,  $p = 0,07$ ), probablemente por la experiencia en valorar la mucosa y en utilizar las clasificaciones endoscópicas. No disponemos actualmente de estudios en los que se considere  $i2b$  como punto de corte, sin poder establecer comparaciones con bibliografía reciente.

La última agrupación definida fue la de RPQ endoscópica grave, que incluye las subcategorías  $i3-i4$ . En los trabajos publicados en los que incluyeron un mayor porcentaje de vídeos  $i3-i4$  la concordancia fue moderada/buena con un kappa de 0,54 y 0,64<sup>181,182</sup>;

superior a los estudios en los que los vídeos fueron mayoritariamente i2. Sin embargo, aunque en nuestro trabajo la concordancia en la valoración de RPQ grave fue moderada ( $\kappa = 0,43$ ), las cifras fueron inferiores a las obtenidas tanto en los estudios publicados como en las otras dos agrupaciones propuestas ( $\kappa = 0,56$  en el punto de corte i2 y  $\kappa = 0,49$  si el punto de corte es i2b). El motivo de esta discrepancia es que, debido a que prácticamente no hay videos i3, la valoración de gravedad se centró en los vídeos i4, con la confusión que supuso esta categoría con respecto a la subcategoría i2b.

Por tanto, como resumen de los resultados obtenidos en el análisis de variabilidad interobservador, todas las agrupaciones (punto de corte en i2, i2b e i3) ofrecen mejor concordancia que el índice de Rutgeerts modificado y la subdivisión i2a e i2b lo único que aporta es un incremento de la discrepancia y de la variabilidad interobservador.

### 9.3. Formación.

Con la aparición del lector central en el ámbito de la endoscopia en la EII y la posibilidad de crear diferentes paneles de lectura, una de las dudas que surgió fue si realmente la figura del *off-site reader* sería imprescindible o sería suficiente con una formación reglada a gastroenterólogos o endoscopistas no expertos, para que fueran capaces de utilizar el índice de Rutgeerts modificado para puntuar las lesiones y así reducir esa variabilidad.

Tras los resultados obtenidos en nuestro trabajo, la inclusión de un único lector central en el panel de lectura no es imprescindible ya que no mejora de forma sustancial la concordancia del índice de Rutgeerts modificado. Esto, añadido a que no siempre se dispone de lectores expertos que puedan realizar la lectura, hace que la solución más accesible sea formar a lectores no expertos como son los lectores locales en la interpretación y el uso del índice de Rutgeerts.

En los últimos años ha cambiado de forma sustancial el método de aprendizaje pasando de un método clásico del ver-hacer-progresar, sin ningún tipo de *feed-back* ni resolución de problemas, a un método basado en competencias en el que la esfera cognitiva resulta fundamental no sólo para la transferencia de conocimientos al alumno, sino también para la

resolución de dudas o corrección de malos hábitos que surgen durante el aprendizaje, incrementando el nivel de confianza y la autonomía del alumno.

La importancia y las ventajas de este modelo de aprendizaje queda reflejada en los numerosos artículos sobre adquisición de competencias que han sido publicados en los últimos años. Sin embargo, la mayoría se basan en competencias puramente técnicas, dejando de lado la esfera cognitiva.

Uno de los objetivos de nuestro trabajo fue valorar precisamente esa esfera cognitiva que incluye la interpretación y clasificación de lesiones endoscópicas propias de la RPQ en la EC. En este punto, se ha comprobado como en la formación hay dos aspectos íntimamente relacionados, por una parte, el **profesor** y por otra parte el **método de evaluación**.

En cuanto al **profesor** se consideró, al igual que en los artículos publicados, a los profesores por su experiencia profesional y conocimiento sobre el tema, no porque dispusieran de un título que acreditara sus aptitudes en este campo. Aunque están surgiendo cursos de formador de formadores<sup>225</sup> para acreditar la actividad docente de estos expertos, todavía no están extendidos a nivel mundial ni tampoco son un requisito obligatorio. Por ello, en la mayoría de los países las personas encargadas de la formación son expertos en la materia sin cursos ni acreditación reglada.

Por otra parte, está el **método de evaluación**. De momento, se desconoce cuál es la forma más adecuada de aprender esa esfera cognitiva y, por tanto, no existen herramientas validadas que nos permitan valorar el método de aprendizaje más adecuado para interpretar las lesiones compatibles con la RPQ endoscópica.

En este estudio se ha equiparado la valoración de lesiones en neóleon mediante vídeos ya grabados con el aprendizaje en la lectura de la cápsula endoscópica. A pesar de no disponer de mucha bibliografía al respecto, recientemente se han publicado varios artículos donde hacen referencia a métodos formativos utilizando precisamente soporte teórico complementado con iconografía mediante vídeos<sup>216, 217</sup>.

Realmente, en la formación no sólo resulta interesante conocer si ésta mejora los datos de concordancia en lectores no expertos, sino si los resultados obtenidos por estos lectores son

capaces de equipararse a los obtenidos por los expertos tanto antes como después del aprendizaje.

Sólo disponemos de dos estudios en los que se analiza la variabilidad entre no expertos sin formación y expertos mostrando diferencias significativas entre ambos grupos a favor de los expertos, pero con grados de concordancia opuestos. En uno de ellos en el que incluyeron a 30 gastroenterólogos generales el índice kappa obtenido fue de 0,67<sup>175</sup> y en el otro estudio que incluyeron a 39 lectores, la concordancia fue inferior con un índice kappa de 0,30 (0,25-0,34)<sup>182</sup>. Probablemente estos resultados tan discordantes y prácticamente opuestos, estén en relación no tanto con el número de vídeos incluidos, que fue similar ya que ambos estudios incluyeron sólo 5 vídeos, sino con la puntuación otorgada a cada uno de ellos. En el primero de ellos no se describió la puntuación de cada lector, pero en el segundo sí, y la puntuación mayoritaria fue i2-i4 (7/10).

En nuestro trabajo, previo a la formación, los MIR obtuvieron una concordancia global débil con un kappa de 0,29, prácticamente la misma que la obtenida en el estudio de Fernandes et al<sup>182</sup> debido a que al igual que en este estudio, en nuestro trabajo la mayoría de puntuaciones otorgadas por los lectores también fueron i2 (7/9). No obstante, estrictamente la diferencia entre expertos y no expertos (kappa = 0,34 vs kappa = 0,29) no fue estadísticamente significativa (p= 0,08).

En base a ambos estudios parece claro que los expertos obtienen una concordancia superior a los no expertos que no han recibido ninguna formación. Por ello, es lógico pensar que una opción podría ser formar a los lectores no expertos que realicen endoscopia en la identificación de las lesiones y utilización de los índices endoscópicos para evitar o al menos minimizar ese grado de discrepancia.

Para valorar la eficacia de la formación, en este trabajo se incluyeron a 9 MIR, tres de ellos en segundo año, otros tres en tercero y los tres restantes en cuarto año de residencia, lo que implicó diferente grado de conocimiento tanto de la EC como de la endoscopia.

La formación administrada supuso una reducción de la variabilidad interobservador tanto global como por año de residencia (Tabla 20, página 138). La concordancia global de los lectores no expertos mejoró pasando de un kappa 0,29 a 0,39 (p = 0,08). Si bien es cierto



que no alcanzó la significación estadística, estaba muy próximo a ella, quedando en un área que nos puede ofrecer cierta incertidumbre. La mejoría fue más evidente al analizar los datos mediante el coeficiente de Kendall, ya que se observó cómo tras la formación los cambios en las puntuaciones no fueron tan extremos, manteniendo un coeficiente por encima de 0,75.

Si en nuestro trabajo comparamos sólo los datos de los MIR en el último año de formación, que estarían más cercanos a los lectores no expertos incluidos en los estudios publicados, los datos de concordancia fueron superiores obteniendo un kappa de 0,39 que pasó a 0,52 tras la formación de forma casi significativa ( $p = 0,06$ ). Esto es debido a que al final de la residencia los MIR presentan una mayor formación en las dos áreas que se imbrican en este trabajo, la endoscopia y la EII, y a que, por tanto, el aprendizaje en este momento es claramente más beneficioso que si se realiza en los primeros años de formación.

A pesar de obtener un aumento de concordancia tras el periodo formativo, tanto de forma global como en los MIR de último año, la mejoría no fue tan espectacular como la obtenida en el estudio de Daperno<sup>216</sup>, que pasó de una concordancia moderada pre-formación con un índice kappa de 0,45 (IC 95% 0,40-0,50), a una concordancia buena-muy buena tras la formación con un índice kappa de 0,79 (IC 95% 0,74-0,83), de forma estadísticamente significativa ( $p = 0,0001$ ).

La buena concordancia obtenida en este trabajo<sup>216</sup>, a diferencia de nuestro estudio, no radica en el propio sistema de formación, ya que en ambos se utilizó el mismo esquema, sino en otros tres aspectos fundamentales. En primer lugar, los lectores incluidos. En el estudio de Daperno, a pesar de que no son expertos en la utilización de los índices endoscópicos, son gastroenterólogos generales con años de experiencia. Sin embargo, en nuestro trabajo son MIR en periodo formativo con diferente grado de conocimientos y de experiencia tanto en la endoscopia como en la EII. En segundo lugar, fue un único experto el que explicó las discrepancias de los vídeos. Por tanto, la lectura posterior que realizaron todos los lectores fue en base a sus criterios, siendo más fácil que todos los alumnos puntuaran de la misma forma. Además, la lectura post-formación tuvo lugar en menos de 24 horas de haber recibido la sesión formativa, a diferencia de nuestro trabajo en el que se realizó pasado dos meses, lo que pudo influir en la eficacia de la formación. Por último, el

número de vídeos incluidos. Sólo incluyeron 5-6 vídeos de pacientes con EC intervenida sin quedar reflejada la puntuación de cada vídeo, a diferencia de nuestro trabajo en el que no sólo se incluyeron más vídeos, sino que la mayoría de ellos (40%) se puntuaron como i2, que a su vez se subdividió en dos categorías, lo que implicó un incremento de esa variabilidad y reducción de la concordancia.

Pero probablemente el dato más importante no sea ese incremento en la concordancia, ya que se mantiene en una concordancia débil, sino que, tras realizar la formación, los lectores no expertos con un kappa de 0,39, se equipararon a los expertos, con un kappa 0,34 ( $p = 0,66$ ). Los lectores no expertos obtuvieron una concordancia discretamente superior, confirmando la eficacia de la formación reglada en la interpretación del índice de Rutgeerts modificado. Concretamente, los MIR en último año de formación obtuvieron una concordancia moderada con un índice kappa de 0,52, superior a la del grupo global de expertos, kappa = 0,34. Este dato demuestra cómo los lectores locales, aunque no sean expertos, tras una formación adecuada son capaces de igualarse a los expertos, constituyendo una alternativa clara en los paneles de lectura a los expertos.

No obstante, al igual que ocurre con los expertos, en la práctica clínica lo importante no es una puntuación concreta sino el punto de corte a partir del cual se considera recurrencia de la enfermedad. En los lectores no expertos, la concordancia pre-formación en las tres agrupaciones, con un índice kappa de 0,54, 0,43 y 0,34 en la primera, segunda y tercera respectivamente, también fue superior a la del índice de Rutgeerts modificado (kappa de 0,29). Sin embargo, los datos obtenidos fueron muy similares a los obtenidos por el grupo de expertos excepto en la valoración de recurrencia grave debido a la dificultad en la puntuación de los vídeos i4 ( tabla 24, página 147). En esta agrupación ( punto de corte i3), los expertos obtuvieron una concordancia superior de forma estadísticamente significativa (kappa 0,43 vs 0,34,  $p=0,007$ ). Tras la formación se produjo un incremento en la concordancia de las tres agrupaciones con un índice kappa de 0,62, 0,54 y 0,50 en la primera, segunda y tercera agrupación respectivamente, superando a los propios expertos y eliminando esas diferencias (tabla 25, página 148). La mejoría más llamativa e incluso significativa se obtuvo con el punto de corte i2b e i3. En el punto de corte i2b la concordancia pasó de un kappa 0,43 a 0,54 ( $p = 0,02$ ) y en el punto de corte i3, el índice

kappa pasó de 0,34 a 0,50 ( $p = 0,007$ ). Estos datos indican que la identificación y valoración de las lesiones que incluyen esas categorías son las que más se benefician de una formación. Debido a ello, se puede plantear que los expertos reciban una formación en la valoración y clasificación de las lesiones, aunque quedaría en debate quien sería el lector encargado de proporcionar esa formación y resolver las dudas entre expertos.

Sólo hay un estudio publicado que compara el índice de Rutgeerts agrupado de los lectores no expertos con los expertos previo a la formación<sup>182</sup>. Al comparar nuestro trabajo con este estudio se observa cómo obtenemos una concordancia en el punto de corte i2 superior (kappa 0,54 vs 0,39). Pero si el objetivo es valorar la recurrencia grave ocurre al contrario, la concordancia de nuestro trabajo es inferior con un kappa 0,34 vs 0,52. El motivo de esta discrepancia son las categorías de los vídeos incluidos. En nuestro trabajo, en la primera agrupación las categorías i2-i4 se valoran en conjunto por ello la concordancia obtenida fue superior a las otras dos agrupaciones. En la recurrencia grave se obtienen datos inferiores debido al mayor número de vídeos i4 y las dudas surgidas en su puntuación con respecto a la categoría i2b.

De momento no disponemos de estudios en los que se compare a lectores expertos con no expertos después de haber sido formados, por lo que desconocemos si nuestros resultados son extrapolables.

Como conclusión de este apartado, la formación supone un aumento de la concordancia tanto del índice de Rutgeerts modificado como del agrupado, resultando estadísticamente significativa en algunos casos como es la determinación de recurrencia grave. Además, tras la formación, los lectores no expertos se igualan a los expertos en términos de concordancia. Sólo con estos datos no se puede afirmar que los lectores no expertos, como son los lectores locales, sean equiparables a los expertos o lectores centrales por varios motivos. En primer lugar, debido a que sólo se ha valorado a los lectores no expertos en una ocasión, a los dos meses de la formación, pero sería necesario reevaluarlos al cabo de más tiempo para determinar si esa mejoría se mantiene en el tiempo. Y, por otra parte, resulta muy complicado conseguir que un lector local sea ciego a la información del paciente, para asegurar esa objetividad en la lectura.

#### 9.4 Lector central.

En la puntuación de los vídeos para valorar la RPQ endoscópica, no existe un patrón oro con el que comparar los diferentes resultados y poder afirmar la existencia de aciertos o fallos. Sólo se disponen de las puntuaciones de diferentes lectores y la variabilidad entre ellos. En este trabajo, además de incluir a un grupo de expertos, se incluyó a otro experto catalogado como lector central. Su elección no fue aleatoria, ya que se trataba de un experto en endoscopia digestiva y concretamente en EII con amplia experiencia en la utilización del índice de Rutgeerts modificado e incluso con participación en ensayos clínicos para valorar RPQ endoscópica. Estos criterios de selección fueron similares a los de otros estudios publicados.<sup>175</sup> En este estudio, al no disponer de un amplio número de lectores, se escogió un único lector central.

El objetivo de incluir a este lector fue conocer la variabilidad de su lectura comparada con el grupo catalogado como expertos, para así valorar si la introducción de un lector central en este ámbito mejora la concordancia del índice y podría ser una solución al problema de variabilidad que tiene *per se*.

Según los datos obtenidos en nuestro trabajo, la inclusión del lector central no mejoró la concordancia prácticamente en ningún caso. La concordancia del índice de Rutgeerts modificado entre el grupo global de expertos y el lector central fue de 0,42, muy similar a la obtenida sin incluirlo, kappa 0,34 ( $p = 0,83$ ) con coeficientes de correlación de Kendall prácticamente idénticos y sin encontrar diferencias entre ambos grupos de expertos (endoscopistas 0,35 vs expertos en EII 0,45  $p = 0,14$ ), homogenizando los resultados obtenidos.

Al analizarlo por agrupaciones, los resultados también fueron equiparables al grupo global de expertos sin encontrar diferencias entre ambos (Tabla 19, página 132). Al comparar el porcentaje de acuerdo entre los lectores expertos y el lector central, se observó cómo la segunda agrupación (i2b) es la que obtuvo peor concordancia con cifras de 73,6% (IC 95% 69,4-77,90) en comparación con las otras dos agrupaciones, de forma estadísticamente significativa ( $p = 0,0012$ ). De hecho, al igual que ocurre con los expertos, la concordancia de la subcategoría i2b fue la que peores resultados ofreció con un kappa de 0,14 siendo i0 la

única subcategoría con una concordancia adecuada ( $\kappa = 0,75$ ). Esto confirma que la profundidad de las úlceras y la afectación de la mucosa subyacente son el origen de la variabilidad interobservador y el motivo de clasificar las lesiones de una forma diferente.

Y, por último, en el apartado de la formación, al igual que en el caso de los lectores expertos, la inclusión de un lector central no modificó los datos de concordancia. En el índice de Rutgeerts modificado se obtuvo una concordancia pre-formación pobre con un  $\kappa$  de 0,32 que mejoró discretamente tras la formación obteniendo un índice  $\kappa$  de 0,40. Sin embargo, los datos son muy similares a los obtenidos sin incluir a este lector ( $\kappa$  pre-formación 0,29 y post-formación de 0,39).

En el índice agrupado, en la única agrupación en la que se observó una concordancia discretamente superior al incluir al lector central fue en la primera, el punto de corte i2. No obstante en ambos casos, con o sin el lector central, la concordancia pre-formación fue moderada, que pasó a ser buena tras la misma. Al incluirlo se obtuvo un  $\kappa$  de 0,60 pre-formación que pasó a 0,74 tras la misma, en comparación a un  $\kappa$  de 0,54 que pasó a 0,62 sin incluirlo. En las otras dos agrupaciones, los datos obtenidos fueron muy similares. En el punto de corte i2b se puede concluir que la concordancia fue moderada en ambos casos. Al incluir al lector central la concordancia pre-formación fue 0,41 que pasó a 0,60 tras la formación, y al no incluirlo fue de 0,44 pre-formación y 0,54 post-formación. En la RPQ grave, la concordancia pre-formación obtenida fue discretamente superior al incluir al lector central ( $\kappa$  0,47 vs 0,34) pero las cifras se igualaron tras la formación con un  $\kappa$  de 0,56 y 0,50 respectivamente (tablas 22 y 23, páginas 132 y 133). De momento, no disponemos de estudios para poder comparar estos datos.

Por tanto, la inclusión del lector central no ofreció ningún beneficio en mejorar la concordancia global del índice de Rutgeerts modificado ni del agrupado ni en los lectores expertos ni en los no expertos, independientemente de su formación. Probablemente esta ausencia de mejoría se deba a que sólo se incluyó un lector central. De momento no disponemos de suficiente información para incluir un número concreto de lectores centrales, pero probablemente uno no sea lo más adecuado por la propia variabilidad del índice e incluso la variabilidad intraobservador, como hemos comentado en el apartado anterior. Aun así, resulta llamativo cómo el lector central siendo endoscopista presenta una

concordancia discretamente superior con el grupo de expertos en EII que con los expertos en endoscopias. Esto explica que quizás no es tan importante el tipo de experto incluido en el panel de lectura y por tanto su experiencia en práctica clínica, sino las variables que componen el índice y que son las que por los motivos que hemos comentado previamente, van a ser las responsables de esa variabilidad.

La ventaja de incluir a lectores centrales es que son lectores que se mantienen actualizados y formados en la lectura de vídeos y en el uso de índices endoscópicos, pero esto no es suficiente y debería ir acompañado de datos de concordancia intraobservador adecuados.

## **10. LIMITACIONES**

---





Algunas de las limitaciones de este trabajo fueron las siguientes:

1. La diferencia en el número de lectores expertos de cada grupo, tres endoscopistas frente a seis expertos clínicos en EII y sólo un lector incluido como lector central.
2. Se repitieron pocos vídeos y además, no están incluidas todas las subcategorías del índice de Rutgeerts modificado. Debido a ello, los datos de variabilidad intraobservador son limitados y hay que interpretarlos con cautela.
3. El número de vídeos catalogados como i3 es escaso, por lo que la variabilidad de esta categoría no fue valorable y la enfermedad grave se centró exclusivamente en los videos catalogados como i4.
4. Las limitaciones en el apartado de formación:
  - No existe ni un programa formativo reglado ni tampoco herramientas validadas para guiar el aprendizaje y realizar una posterior evaluación del mismo.  
A pesar de ello, se ha utilizado en mismo método formativo que en el resto de trabajos, un soporte iconográfico y posterior discusión de aspectos controvertidos hasta llegar a un consenso.
  - Los profesores encargados de la formación de los lectores no expertos son expertos en la materia, pero no han realizado cursos o disponen de acreditación que certifique que son profesionales aptos para la formación. De momento, ya que en nuestro país no disponemos ni de herramientas ni de un sistema general de formación acreditado en la endoscopia digestiva, la formación en este ámbito sigue quedando en manos de profesionales expertos en la materia. Realmente estas dos limitaciones no son propias en exclusiva de nuestro trabajo sino comunes a todos los trabajos realizados hasta el momento e implican de por sí cierta variabilidad en los estudios.
  - El tiempo transcurrido hasta la lectura postformación. Se dejó transcurrir dos meses tras la formación para reevaluar a los lectores no expertos. Quizás lo interesante sería reevaluar de nuevo a los lectores pasado un año o más del aprendizaje para ver si los resultados se mantienen en el tiempo.



## **11. CONCLUSIONES**

---



1. En nuestro estudio la reproducibilidad del índice de Rutgeerts modificado es baja, por ello sería conveniente replantear su uso en la valoración de la RPQ endoscópica.
2. Los resultados obtenidos sugieren que las categorías i2a e i2b deberían valorarse en conjunto, ya que hacerlo de forma separada sólo aporta un incremento de la variabilidad y reducción de la concordancia entre los lectores.
3. La agrupación de categorías del índice de Rutgeerts modificado mejora la concordancia, sobre todo si el punto de corte se establece en i2.
4. La inclusión de un lector central no incrementa la concordancia. No obstante, se puede plantear su inclusión para mantener una estabilidad en los criterios de clasificación, una formación y actualización constante en la lectura de vídeos y el uso de índices endoscópicos, así como asegurar el ciego a la información clínica del paciente.
5. La formación a los lectores no expertos incrementa de forma notable la concordancia tanto del índice de Rutgeerts modificado como de las tres agrupaciones propuestas, alcanzando una concordancia similar o incluso superior a la del grupo de expertos. Debido a estos resultados, se debería considerar la formación continuada en los lectores expertos.



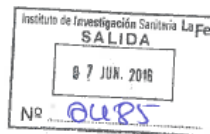
## **12. ANEXOS**

---





## Anexo 1: Documento de aprobación.



Dra. Noelia Alonso Lázaro  
Servicio Medicina Digestiva

**Asunto:** Autorización Inicio Estudio.

Valencia, 24 de Mayo de 2016

Adjunto le remito copia de los Informes Científico y Ético de Investigación, en el que se acuerda informar **favorablemente** a las modificaciones solicitadas sobre el Proyecto de Investigación titulado "VARIABILIDAD INTEROBSERVADOR EN EL ÍNDICE DE RUTGEERTS PARA DETERMINAR LA RECURRENCIA POSTQUIRÚRGICA POR COLONOSCOPIA EN LA ENFERMEDAD DE CROHN", por usted presentado.

A la vista de los dictámenes emitidos, dicho Proyecto, puede iniciarse y llevarse a cabo.

Atentamente,



José María Millán Salvador  
Director General

## Anexo 2: Formulario de lectura de los vídeos.

Se incluye el formulario 4 como ejemplo de los formularios elaborados en la plataforma Formularios de Google.

**RECURRENCIA POSTQUIRÚRGICA EN LA ENFERMEDAD DE CROHN. ANÁLISIS DE VARIABILIDAD INTEROBSERVADOR DEL ÍNDICE DE RUTGEERTS.**

\* Required

1. **Nombre y apellidos \***

\_\_\_\_\_

**FASE 4**  
Responda a las preguntas empleando el índice de Rutgeerts

2. **Paciente 34 \***  
VÍDEO: <https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/wi3o3d9a7H2KQih>  
Mark only one oval.

0

1

2a

2b

3

4

3. **Calidad del vídeo \***  
Mark only one oval.

Adecuada

Inadecuada

**FASE 4**  
Responda a las preguntas empleando el índice de Rutgeerts

4.

**Paciente 35 \***

VÍDEO:

[https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/xbP8D2SII\\_vLWQc8](https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/xbP8D2SII_vLWQc8)

Mark only one oval.

- 0
- 1
- 2a
- 2b
- 3
- 4

5.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada
- Inadecuada

## FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

6.

**Paciente 36 \***

VÍDEO:

<https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/qvblKkN8XCc0HSH>

Mark only one oval.

- 0
- 1
- 2a
- 2b
- 3
- 4

7.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada
- Inadecuada

## FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

8.

**Paciente 37 \***

VÍDEO: <https://aplicaciones.lafe.san.qva.es/cloudlafe/index.php/s/v2ikZGvPNDNxO0p>

Mark only one oval.

- 0  
 1  
 2a  
 2b  
 3  
 4

9.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada  
 Inadecuada

#### FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

10.

**Paciente 38 \***

VÍDEO:

<https://aplicaciones.lafe.san.qva.es/cloudlafe/index.php/s/DoYo5EAo7RRInEE>

Mark only one oval.

- 0  
 1  
 2a  
 2b  
 3  
 4

11.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada  
 Inadecuada

#### FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

12.

**Paciente 39 \***

VÍDEO: <https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/1fORnshmpUZIP03>

Mark only one oval.

- 0
- 1
- 2a
- 2b
- 3
- 4

13.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada
- Inadecuada

#### FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

14.

**Paciente 40 \***

VÍDEO: <https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/SALoJfpxBdVStBy>

Mark only one oval.

- 0
- 1
- 2a
- 2b
- 3
- 4

15.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada
- Inadecuada

#### FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

16.

**Paciente 41 \***

VÍDEO: <https://aplicaciones.lafe.san.gva.es/cloudlafa/index.php/s/unudQoK9y0ih6ST>

Mark only one oval.

0

1

2a

2b

3

4

17.

**Calidad del video \***

Mark only one oval.

Adecuada

Inadecuada

## FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

18.

**Paciente 42 \***

VÍDEO: <https://aplicaciones.lafe.san.gva.es/cloudlafa/index.php/s/ZD3VW3JPvQ59LuD>

Mark only one oval.

0

1

2a

2b

3

4

19.

**Calidad del video \***

Mark only one oval.

Adecuada

Inadecuada

## FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

20.

**Paciente 43 \***

VÍDEO: <https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/IY4u7s8kt5ZxvxC>

Mark only one oval.

- 0
- 1
- 2a
- 2b
- 3
- 4

21.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada
- Inadecuada

## FASE 4

Responda a las preguntas empleando el índice de Rutgeerts

22.

**Paciente 44 \***

VÍDEO:

<https://aplicaciones.lafe.san.gva.es/cloudlafe/index.php/s/hUYCoZhVxXVQjBP>

Mark only one oval.

- 0
- 1
- 2a
- 2b
- 3
- 4

23.

**Calidad del vídeo \***

Mark only one oval.

- Adecuada
- Inadecuada

### Anexo 3: Hoja basal de recogida de datos.

#### DATOS PERSONALES

Número de paciente incluido:

Fecha de nacimiento:

Sexo:  Mujer  Hombre

Raza:  Caucásica  Africana  Asiática  Otras

Fumador:  No  Sí  Exfumador

Alcohol:  No  Sí

#### HISTORIA ENFERMEDAD DE CROHN

Año diagnóstico de EC:

Patrón:  inflamatorio  estenosante  perforante  NA

Localización:  íleon (L1)  Íleon más colon (L3)

Perianal:  No  Sí

Afectación de tracto superior (L4):  No  Sí

Número de intervenciones:

Fecha de intervenciones:

Motivo de la última intervención:

- Actividad/plastrón/refractario tratamiento
- Estenosis/cuadro suboclusivo
- Fístulas

Tipo de intervención:  resección ileocecal  hemicolectomía derecha

Tipo de anastomosis:  laterolateral  terminolateral  terminoterminal

Complicaciones:  No  Sí → tipo:

Longitud resección intestinal (cm):



Tratamiento como prevención de la recurrencia:  No  Sí.

Tipo de tratamiento preventivo:

- ATB
- Corticoides
- AZA/6-MP
- IFX
- ADA
- IFX+ADA/IFX+AZA
- Otro

#### Anexo 4: Variabilidad intraobservador.

##### Grupo Endoscopia Digestiva

	Respuesta	kappa	SE	Z	p
Lector 1	0	1	0,408248	2,44949	0,0072
	1	1	0,408248	2,44949	0,0072
	2a	1	0,408248	2,44949	0,0072
	2b	1	0,408248	2,44949	0,0072
	General	1	0,25	4	0,0000
Lector 2	0	1	0,408248	2,44949	0,0072
	1	1	0,408248	2,44949	0,0072
	2a	1	0,408248	2,44949	0,0072
	2b	1	0,408248	2,44949	0,0072
	General	1	0,25	4	0,0000
Lector 3	0	1	0,408248	2,44949	0,0072
	1	1	0,408248	2,44949	0,0072
	2a	1	0,408248	2,44949	0,0072
	2b	1	0,408248	2,44949	0,0072
	General	1	0,25	4	0,0000

Grupo de expertos en Enfermedad Inflamatoria Intestinal.

	Respuesta	kappa	SE	Z	p
Lector 1	0	0,555556	0,408248	1,36083	0,0868
	1	1	0,408248	2,44949	0,0072
	2a	0	0,408248	0	0,6879
	2b	0	0,408248	0	0,5881
	3	*	*	*	*
	4	1	0,408248	2,44949	0,0072
	General	0,56364	0,218308	2,58184	0,0049
Lector 2	0	1	0,408248	2,44949	0,0072
	1	1	0,408248	2,44949	0,0072
	2a	1	0,408248	2,44949	0,0072
	2b	1	0,408248	2,44949	0,0072
	3	*	*	*	*
	4	1	0,408248	2,44949	0,0072
	General	1	0,210280	4,75556	0,0000
Lector 3	0	1	0,408248	2,44949	0,0072
	1	1	0,408248	2,44949	0,0072
	2a	1	0,408248	2,44949	0,0072
	2b	1	0,408248	2,44949	0,0072
	3	1	0,408248	2,44949	0,0072
	4	*	*	*	*
	General	1	0,210280	4,75556	0,0000
Lector 4	0	*	*	*	*
	1	0,65714	0,408248	1,60966	0,0537
	2a	*	*	*	*
	2b	1	0,408248	2,44949	0,0072
	3	1	0,408248	2,44949	0,007
	4	0	0,408248	0	0,5881
	General	0,75510	0,263073	2,87032	0,0021

	Respuesta	kappa	SE	Z	p
Lector 5	0	1	0,408248	2,44949	0,0072
	1	1	0,408248	2,44949	0,0072
	2a	*	*	*	*
	2b	*	*	*	*
	3	1	0,408248	2,44949	0,0072
	4	1	0,40824	2,44949	0,0072
	General	1	0,245271	4,07713	0,0000

### Anexo 5: Resultados con un experto en EII como lector central.

<b>Global</b>	% acuerdo e IC 95%	kappa	p	Coeficiente Kendall
Entre evaluadores	10,91 (4,11-22,25)	0,366439	0,000	0,7727
Comparado con LC	7,27 (2,02-17,59)	0,318049	0,0000	0,572771
<b>Endoscopias</b>				
Entre evaluadores	29,09 (17,63-42,9)	0,414762	0,0000	0,8147
Comparado con LC	20,00 (10,43-32,97)	0,338356	0,0000	0,575737
<b>EII</b>				
Entre evaluadores	20,00 (10,43-32,97)	0,325511	0,0000	0,799176
Comparado con LC	14,55 (6,5-26,66)	0,297743	0,0000	0,569804



## **13. BIBLIOGRAFÍA**

---





1. Baumgart, D. C. & Sandborn, W. J. Inflammatory bowel disease: clinical aspects and established and evolving therapies. *Lancet* **369**, 1641–1657 (2007).
2. Burisch, J., Jess, T., Martinato, M. & Lakatos, P. L. The burden of inflammatory bowel disease in Europe. *J. Crohn's Colitis* **7**, 322–337 (2013).
3. Molodecky, N. A. *et al.* Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* **142**, 46–54.e42 (2012).
4. M'Koma, A. E. Inflammatory bowel disease: An expanding global health problem. *Clin. Med. Insights Gastroenterol.* **6**, 33–47 (2013).
5. Burisch, J. *et al.* East-West gradient in the incidence of inflammatory bowel disease in Europe: The ECCO-EpiCom inception cohort. *Gut* **63**, 588–597 (2014).
6. Leddin, D., Tamim, H. & Levy, A. R. Decreasing incidence of inflammatory bowel disease in Eastern Canada: A population database study. *BMC Gastroenterol.* **14**, 1–8 (2014).
7. Ng, S. C. *et al.* Incidence and phenotype of inflammatory bowel disease based on results from the Asia-Pacific Crohn's and colitis epidemiology study. *Gastroenterology* **145**, 158–165 (2013).
8. Thia, K. T., Loftus, E. V., Sandborn, W. J. & Yang, S. K. An update on the epidemiology of inflammatory bowel disease in Asia. *Am. J. Gastroenterol.* **103**, 3167–3182 (2008).
9. Ng, S. C. *et al.* Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* **390**, (2017).
10. Gilmore, J. H. NIH Public Access. *North* **29**, 1883–1889 (2008).
11. Baumgart, D. C. & Carding, S. R. Inflammatory bowel disease: cause and immunobiology. *Lancet* **369**, 1627–1640 (2007).
12. Hamilton, M. J. & Blumberg, R. S. Their Therapeutic Relevance. **47**, 1–8 (2015).
13. Harrison, O. J. & Maloy, K. J. Innate immune activation in intestinal homeostasis. *J. Innate Immun.* **3**, 585–593 (2011).
14. Naser, S. A. *et al.* Role of ATG16L, NOD2 and IL23R in Crohn's disease pathogenesis. *World J. Gastroenterol.* **18**, 412–424 (2012).
15. Newburger, Jane W. Takashi, Masato. Alexa Steven. Colan, Elise. Duffy, Fulton, D. *et al.* The New England Journal of Medicine Downloaded from nejm.org at MOUNT SINAI SCHOOL OF MEDICINE on January 15, 2014. For personal use only. No other uses without permission. Copyright © 1991 Massachusetts Medical Society.

All rights reserved. (1991).

16. Lowe, A. M. *et al.* Epidemiology of Crohn's disease in Québec, Canada. *Inflamm. Bowel Dis.* (2009). doi:10.1002/ibd.20756
17. Abraham, C. & Cho, J. H. Inflammatory Bowel Disease. *New Engl. J. Med.* **361**, 2066–2078 (2009).
18. van der Sloot, K. W. J., Amini, M., Peters, V., Dijkstra, G. & Alizadeh, B. Z. Inflammatory Bowel Diseases. *Inflamm. Bowel Dis.* **23**, 1 (2017).
19. Somerville, K. W., Logan, R. F., Edmond, M. & Langman, M. J. Smoking and Crohn's disease. *Br. Med. J. (Clin. Res. Ed)*. **289**, 954–6 (1984).
20. Barua, R. S. *et al.* Dysfunctional endothelial nitric oxide biosynthesis in healthy smokers with impaired endothelium-dependent vasodilatation. *Circulation* **104**, 1905–1910 (2001).
21. Johnson, G. J., Cosnes, J. & Mansfield, J. C. Review article: Smoking cessation as primary therapy to modify the course of Crohn's disease. *Alimentary Pharmacology and Therapeutics* **21**, 921–931 (2005).
22. Stępień, E., Miszański-Jamka, T., Kapusta, P., Tylko, G. & Pasowicz, M. Beneficial effect of cigarette smoking cessation on fibrin clot properties. *J. Thromb. Thrombolysis* **32**, 177–182 (2011).
23. Wakefield, A. J. *et al.* PATHOGENESIS OF CROHN'S DISEASE: MULTIFOCAL GASTROINTESTINAL INFARCTION. *Lancet* **334**, 1057–1062 (1989).
24. Higuchi, L. M. *et al.* A Prospective Study of Cigarette Smoking and the Risk of Inflammatory Bowel Disease in Women. *Am. J. Gastroenterol.* **107**, 1399–1406 (2012).
25. Cosnes, J., Carbonnel, F., Beaugerie, L., Le Quintrec, Y. & Gendre, J. Effects of cigarette smoking on the long-term course of Crohn's disease. *Gastroenterology* **110**, 424–431 (1996).
26. Nunes, T. *et al.* Smoking does influence disease behaviour and impacts the need for therapy in Crohn's disease in the biologic era. *Aliment. Pharmacol. Ther.* **38**, 752–760 (2013).
27. Louis, E. *et al.* Early development of stricturing or penetrating pattern in Crohn's disease is influenced by disease location, number of flares, and smoking but not by NOD2/CARD15 genotype. *Gut* **52**, 552–557 (2003).
28. Kuenzig, M. E. *et al.* Smoking influences the need for surgery in patients with the inflammatory bowel diseases: A systematic review and meta-analysis incorporating disease duration. *BMC Gastroenterol.* **16**, (2016).
29. To, N., Gracie, D. J. & Ford, A. C. Systematic review with meta-analysis: the

- adverse effects of tobacco smoking on the natural history of Crohn's disease. *Aliment. Pharmacol. Ther.* **43**, 549–561 (2016).
30. Biedermann, L. *et al.* Smoking cessation alters intestinal microbiota: Insights from quantitative investigations on human fecal samples using FISH. *Inflamm. Bowel Dis.* **20**, 1496–1501 (2014).
  31. Ananthakrishnan, A. N. *et al.* A prospective study of long-term intake of dietary fiber and risk of Crohn's disease and ulcerative colitis. *Gastroenterology* **145**, 970–977 (2013).
  32. Opstelten, J. L. *et al.* Dairy Products, Dietary Calcium, and Risk of Inflammatory Bowel Disease: Results From a European Prospective Cohort Investigation. *Inflamm. Bowel Dis.* **22**, 1403–11 (2016).
  33. Hlavaty, T., Krajcovicova, A. & Payer, J. Vitamin D therapy in inflammatory bowel diseases: who, in what form, and how much? *J. Crohns. Colitis* **9**, 198–209 (2015).
  34. Ananthakrishnan, A. N. *et al.* Higher predicted vitamin D status is associated with reduced risk of crohn's disease. *Gastroenterology* **142**, 482–489 (2012).
  35. Hou, J. K., Abraham, B. & El-Serag, H. Dietary Intake and Risk of Developing Inflammatory Bowel Disease: A Systematic Review of the Literature. *Am. J. Gastroenterol.* **106**, 563–573 (2011).
  36. Ananthakrishnan, A. N. *et al.* Long-term intake of dietary fat and risk of ulcerative colitis and Crohn's disease. *Gut* **63**, 776–784 (2014).
  37. Uchiyama, K. *et al.* N-3 polyunsaturated fatty acid diet therapy for patients with inflammatory bowel disease. *Inflamm. Bowel Dis.* **16**, 1696–1707 (2010).
  38. Lev-Tzion, R., Griffiths, A. M., Leder, O. & Turner, D. Omega 3 fatty acids (fish oil) for maintenance of remission in Crohn's disease. *Cochrane database Syst. Rev.* **2**, CD006320 (2014).
  39. D'Argenio, V. *et al.* An altered gut microbiome profile in a child affected by Crohn's disease normalized after nutritional therapy. *American Journal of Gastroenterology* **108**, 851–852 (2013).
  40. H., K. *et al.* Association between Long-term Oral Contraceptive Use and Risk of Crohn's Disease Complications in a Nationwide Study. *Gastroenterology* **150**, 1561–1567.e1 (2016).
  41. Ungaro, R. *et al.* Antibiotics Associated With Increased Risk of New-Onset Crohn's Disease But Not Ulcerative Colitis: A Meta-Analysis. *Am. J. Gastroenterol.* **109**, 1728–1738 (2014).
  42. Sands, B. E. From symptom to diagnosis: Clinical distinctions among various forms of intestinal inflammation. *Gastroenterology* **126**, 1518–1532 (2004).

43. Charpentier, C. *et al.* Natural history of elderly-onset inflammatory bowel disease: A population-based cohort study. *Gut* **63**, 423–432 (2014).
44. Halpin, S. J. & Ford, A. C. Prevalence of symptoms meeting criteria for irritable bowel syndrome in inflammatory bowel disease: Systematic review and meta-analysis. *American Journal of Gastroenterology* **107**, 1474–1482 (2012).
45. Mücke, V., Mücke, M. M., Raine, T. & Bettenworth, D. Diagnosis and treatment of anemia in patients with inflammatory bowel disease. *Ann. Gastroenterol. Ann Gastroenterol* **30**, 15–22 (2017).
46. Lennard-Jones, J. E. & Shivananda, S. Clinical uniformity of inflammatory bowel disease a presentation and during the first year of disease in the north and south of Europe. EC-IBD Study Group. *Eur J Gastroenterol Hepatol* **9**, 353–359 (1997).
47. Peyrin-Biroulet, L. *et al.* Perianal Crohn’s disease findings other than fistulas in a population-based cohort. *Inflamm. Bowel Dis.* **18**, 43–48 (2012).
48. Harbord, M. *et al.* The first european evidence-based consensus on extra-intestinal manifestations in inflammatory bowel disease. *J. Crohn’s Colitis* **10**, 239–254 (2016).
49. Bernstein, C. N., Blanchard, J. F., Rawsthorne, P. & Yu, N. The prevalence of extraintestinal diseases in inflammatory bowel disease: a population-based study. *Am. J. Gastroenterol.* **96**, 1116–1122 (2001).
50. Hagen, J. W., Swoger, J. M. & Grandinetti, L. M. Cutaneous Manifestations of Crohn Disease. *Dermatologic Clinics* **33**, 417–431 (2015).
51. Thomas, A. S. & Lin, P. Ocular manifestations of inflammatory bowel disease. *Current Opinion in Ophthalmology* **27**, 552–560 (2016).
52. Yarur, A. J., Czul, F. & Levy, C. Hepatobiliary manifestations of inflammatory bowel disease. *Inflammatory Bowel Diseases* **20**, 1655–1667 (2014).
53. Gizard, E., Ford, A. C., Bronowicki, J.-P. & Peyrin-Biroulet, L. Systematic review: the epidemiology of the hepatobiliary manifestations in patients with inflammatory bowel disease. *Aliment. Pharmacol. Ther.* **40**, 3–15 (2014).
54. Reinshagen, M. Osteoporosis in inflammatory bowel disease. *Journal of Crohn’s and Colitis* **2**, 202–207 (2008).
55. Nguyen, G. C. & Sam, J. Rising prevalence of venous thromboembolism and its impact on mortality among hospitalized inflammatory bowel disease patients. *Am. J. Gastroenterol.* **103**, 2272–2280 (2008).
56. Van Assche, G. *et al.* The second European evidence-based Consensus on the diagnosis and management of Crohn’s disease: Definitions and diagnosis. *J. Crohns. Colitis* **4**, 7–27 (2010).

57. Lennard-Jones, J. E. Classification of inflammatory bowel disease. *Scand. J. Gastroenterol. Suppl.* **170**, 2–6; discussion 16-9 (1989).
58. Magro, F. *et al.* European consensus on the histopathology of inflammatory bowel disease. *J. Crohn's Colitis* **7**, 827–851 (2013).
59. Reinisch, S. *et al.* An index with improved diagnostic accuracy for the diagnosis of Crohn's disease derived from the Lennard-Jones criteria. *Aliment. Pharmacol. Ther.* **44**, 601–611 (2016).
60. Farmer, R. G., Hawk, W. A. & Turnbull, R. B. Clinical Patterns in Crohn's Disease: A Statistical Study of 615 Cases. *Gastroenterology* **68**, 627–635 (1975).
61. Greenstein, a J. *et al.* Perforating and non-perforating indications for repeated operations in Crohn's disease: evidence for two clinical forms. *Gut* **29**, 588–592 (1988).
62. Gasche, C. *et al.* A simple classification of Crohn's disease: Report of the working party for the world congresses of gastroenterology, Vienna 1998. *Inflamm. Bowel Dis.* **6**, 8–15 (2000).
63. Silverberg, M. S. *et al.* Toward an Integrated Clinical, Molecular and Serological Classification of Inflammatory Bowel Disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can. J. Gastroenterol.* **19**, 5A-36A (2005).
64. Nuij, V. J. A. A. *et al.* Phenotype of inflammatory bowel disease at diagnosis in the Netherlands: A population-based inception cohort study (the Delta Cohort). *Inflamm. Bowel Dis.* **19**, 2215–2222 (2013).
65. Sjöberg, D. *et al.* Incidence and natural history of ulcerative colitis in the Uppsala Region of Sweden 2005-2009 - Results from the IBD Cohort of the Uppsala Region (ICURE). *J. Crohn's Colitis* **7**, e351–e357 (2013).
66. Lakatos, L. *et al.* Incidence, disease phenotype at diagnosis, and early disease course in inflammatory bowel diseases in Western Hungary, 2002-2006. *Inflamm. Bowel Dis.* **17**, 2558–2565 (2011).
67. Walsh, C. M. *et al.* Development of the gastrointestinal endoscopy competency assessment tool for pediatric colonoscopy (GiECATKIDS). *J. Pediatr. Gastroenterol. Nutr.* **59**, 480–486 (2014).
68. Ramadas, A. V., Gunesh, S., Thomas, G. A. O., Williams, G. T. & Hawthorne, A. B. Natural history of Crohn's disease in a population-based cohort from Cardiff (1986-2003): a study of changes in medical treatment and surgical resection rates. *Gut* **59**, 1200–1206 (2010).
69. Best, W. R., Bectel, J. M., Singleton, J. W. & Kern, F. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology* **70**, 439–44 (1976).

70. Harvey, R. F. & Bradshaw, J. M. A SIMPLE INDEX OF CROHN'S-DISEASE ACTIVITY. *Lancet* **315**, 514 (1980).
71. Vermeire, S., Schreiber, S., Sandborn, W. J., Dubois, C. & Rutgeerts, P. Correlation Between the Crohn's Disease Activity and Harvey-Bradshaw Indices in Assessing Crohn's Disease Severity. *Clin. Gastroenterol. Hepatol.* **8**, 357–363 (2010).
72. van Hees, P. A., van Elteren, P. H., van Lier, H. J. & van Tongeren, J. H. An index of inflammatory activity in patients with Crohn's disease. *Gut* **21**, 279–86 (1980).
73. Irvine, E. J. Usual therapy improves perianal Crohn's disease as measured by a new disease activity index. McMaster IBD Study Group. *J. Clin. Gastroenterol.* **20**, 27–32 (1995).
74. Panés, J. *et al.* Systematic review: The use of ultrasonography, computed tomography and magnetic resonance imaging for the diagnosis, assessment of activity and abdominal complications of Crohn's disease. *Aliment. Pharmacol. Ther.* **34**, 125–145 (2011).
75. Lee, S. S. *et al.* CT of prominent pericolic or perienteric vasculature in patients with Crohn's disease: Correlation with clinical disease activity and findings on barium studies. *Am. J. Roentgenol.* **179**, 1029–1036 (2002).
76. Greenup, A. J., Bressler, B. & Rosenfeld, G. Medical imaging in small bowel Crohn's disease - Computer tomography enterography, magnetic resonance enterography, and ultrasound: 'Which one is the best for what?'. *Inflamm. Bowel Dis.* **22**, 1246–1261 (2016).
77. Chiorean, M. V. *et al.* Correlation of CT enteroclysis with surgical pathology in Crohn's disease. *Am. J. Gastroenterol.* **102**, 2541–2550 (2007).
78. Rimola, J. *et al.* Magnetic resonance for assessment of disease activity and severity in ileocolonic Crohn's disease. *Gut* **58**, 1113–1120 (2009).
79. Rimola, J. *et al.* Magnetic resonance imaging for evaluation of Crohn's disease: Validation of parameters of severity and quantitative index of activity. *Inflamm. Bowel Dis.* **17**, 1759–1768 (2011).
80. Ordás, I. *et al.* Accuracy of magnetic resonance enterography in assessing response to therapy and mucosal healing in patients with Crohn's disease. *Gastroenterology* **146**, 374-382.e1 (2014).
81. Coimbra, A. J. F. *et al.* Magnetic resonance enterography is feasible and reliable in multicenter clinical trials in patients with Crohn's disease, and may help select subjects with active inflammation. *Aliment. Pharmacol. Ther.* **43**, 61–72 (2016).
82. Steward, M. J. *et al.* Non-perforating small bowel Crohn's disease assessed by MRI enterography: Derivation and histopathological validation of an MR-based activity index. *Eur. J. Radiol.* **81**, 2080–2088 (2012).

83. Prezzi, D. *et al.* Monitoring Crohn's disease during anti-TNF- $\alpha$  therapy: validation of the magnetic resonance enterography global score (MEGS) against a combined clinical reference standard. *Eur. Radiol.* **26**, 2107–2117 (2016).
84. Oussalah, A. *et al.* Diffusion-weighted magnetic resonance without bowel preparation for detecting colonic inflammation in inflammatory bowel disease. *Gut* **59**, 1056–1065 (2010).
85. Pariente, B. *et al.* Development of the Crohn's disease digestive damage score, the Lémann score. *Inflamm. Bowel Dis.* **17**, 1415–1422 (2011).
86. Pariente, B. *et al.* Development of the Lémann index to assess digestive tract damage in patients with Crohn's disease. *Gastroenterology* **148**, 52-63.e3 (2015).
87. S., V., G., V. A. & P., R. C-reactive protein as a marker for inflammatory bowel disease. *Inflamm. Bowel Dis.* **10**, 661–665 (2004).
88. Lehmann, F. S., Burri, E. & Beglinger, C. The role and utility of faecal markers in inflammatory bowel disease. *Therapeutic Advances in Gastroenterology* **8**, 23–36 (2015).
89. D'Haens, G. *et al.* Fecal calprotectin is a surrogate marker for endoscopic lesions in inflammatory bowel disease. *Inflamm. Bowel Dis.* **18**, 2218–2224 (2012).
90. Sipponen, T. *et al.* Faecal calprotectin and lactoferrin are reliable surrogate markers of endoscopic response during Crohn's disease treatment. *Scand. J. Gastroenterol.* **45**, 325–331 (2010).
91. Walsh, C. M. *et al.* Gastrointestinal endoscopy competency assessment tool: Reliability and validity evidence. *Gastrointest. Endosc.* **81**, 1417-1424.e2 (2015).
92. Lin, J.-F. *et al.* Meta-analysis. *Inflamm. Bowel Dis.* **20**, 1407–1415 (2014).
93. Gisbert, J. P. *et al.* Fecal calprotectin and lactoferrin for the prediction of inflammatory bowel disease relapse. *Inflamm. Bowel Dis.* **15**, 1190–1198 (2009).
94. Sipponen, T. *et al.* Crohn's disease activity assessed by fecal calprotectin and lactoferrin: Correlation with Crohn's disease activity index and endoscopic findings. *Inflamm. Bowel Dis.* **14**, 40–46 (2008).
95. Schoepfer, A. M. *et al.* Fecal calprotectin correlates more closely with the simple endoscopic score for crohn's disease (SES-CD) than CRP, blood leukocytes, and the CDAI. *Am. J. Gastroenterol.* **105**, 162–169 (2010).
96. Veloso, F. T., Ferreira, J. T., Barros, L. & Almeida, S. Clinical outcome of Crohn's disease: analysis according to the vienna classification and clinical activity. *Inflamm. Bowel Dis.* (2001). doi:10.1097/00054725-200111000-00005
97. Romberg-Camps, M. J. L. *et al.* Influence of phenotype at diagnosis and of other potential prognostic factors on the course of inflammatory bowel disease. *Am. J.*

- Gastroenterol.* **104**, 371–383 (2009).
98. Peyrin-biroulet, L., Loft, E. V, Colombel, J. & Sandborn, W. J. The Natural History of Adult Crohn ' s Disease in Population-Based Cohorts. *Am. J. Gastroenterol.* **105**, 289–297 (2009).
  99. Frolkis, A. D. *et al.* Risk of surgery for inflammatory bowel diseases has decreased over time: A systematic review and meta-analysis of population-based studies. *Gastroenterology* (2013). doi:10.1053/j.gastro.2013.07.041
  100. Rutgeerts, P. *et al.* Natural history of recurrent Crohn's disease at the ileocolonic anastomosis after curative surgery. *Gut* **25**, 665–672 (1984).
  101. Rutgeerts, P. *et al.* Predictability of the postoperative course of Crohn's disease. *Gastroenterology* **99**, 956–963 (1990).
  102. Buisson, A., Chevaux, J. B., Allen, P. B., Bommelaer, G. & Peyrin-Biroulet, L. Review article: The natural history of postoperative Crohn's disease recurrence. *Alimentary Pharmacology and Therapeutics* (2012). doi:10.1111/j.1365-2036.2012.05002.x
  103. D'Haens, G. R. *et al.* Early lesions of recurrent Crohn's disease caused by infusion of intestinal contents in excluded ileum. *Gastroenterology* (1998). doi:10.1016/S0016-5085(98)70476-7
  104. Lopez, J. *et al.* Natural history of Crohn's disease following total colectomy and end ileostomy. *Inflamm. Bowel Dis.* (2014). doi:10.1097/MIB.0000000000000072
  105. Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci.* (2008). doi:10.1073/pnas.0804812105
  106. Neut, C. *et al.* Changes in the bacterial flora of the neoterminal ileum after ileocolonic resection for Crohn's disease. *Am. J. Gastroenterol.* (2002). doi:10.1016/S0002-9270(02)03969-2
  107. Dey, N., Soergel, D. A. W., Repo, S. & Brenner, S. E. Association of gut microbiota with post-operative clinical course in Crohn's disease. *BMC Gastroenterol.* (2013). doi:10.1186/1471-230X-13-131
  108. Wright, E. K. *et al.* Microbial Factors Associated with Postoperative Crohn's Disease Recurrence. *J. Crohns. Colitis* (2016). doi:10.1093/ecco-jcc/jjw136
  109. De Cruz, P. *et al.* Association between specific mucosa-associated microbiota in Crohn's disease at the time of resection and subsequent disease recurrence: A pilot study. *J. Gastroenterol. Hepatol.* (2015). doi:10.1111/jgh.12694
  110. Torres, J. *et al.* Predicting Outcomes to Optimize Disease Management in Inflammatory Bowel Diseases. *J. Crohn's Colitis* (2016). doi:10.1093/ecco-jcc/jjw116



111. Pascua, M., Su, C., Lewis, J. D., Brensinger, C. & Lichtenstein, G. R. Meta-analysis: factors predicting post-operative recurrence with placebo therapy in patients with Crohn's disease. *Aliment. Pharmacol. Ther.* (2008). doi:10.1111/j.1365-2036.2008.03774.x
112. Reese, G. E. *et al.* The effect of smoking after surgery for Crohn's disease: A meta-analysis of observational studies. *Int. J. Colorectal Dis.* **23**, 1213–1221 (2008).
113. Nunes, T. *et al.* Impact of Smoking Cessation on the Clinical Course of Crohn's Disease Under Current Therapeutic Algorithms: A Multicenter Prospective Study. *Am. J. Gastroenterol.* **111**, 411–419 (2016).
114. Nunes, T. *et al.* High smoking cessation rate in Crohn's disease patients after physician advice - The TABACROHN Study. *J. Crohn's Colitis* **7**, 202–207 (2013).
115. Strong, S. *et al.* Clinical practice guideline for the surgical management of Crohn's disease. *Dis. Colon Rectum* **58**, 1021–1036 (2015).
116. Decousus, S. *et al.* Myenteric plexitis is a risk factor for endoscopic and clinical postoperative recurrence after ileocolonic resection in Crohn's disease. *Dig. Liver Dis.* (2016). doi:10.1016/j.dld.2016.02.023
117. Lemmens, B. *et al.* Submucosal Plexitis as a Predictive Factor for Postoperative Endoscopic Recurrence in Patients with Crohn's Disease Undergoing a Resection with Ileocolonic Anastomosis: Results from a Prospective Single-centre Study. *J. Crohns. Colitis* (2017). doi:10.1093/ecco-jcc/jjw135
118. Annese, V. *et al.* European evidence based consensus for endoscopy in inflammatory bowel disease. *J. Crohn's Colitis* **7**, 982–1018 (2013).
119. Rosa, B., Moreira, M. J., Rebelo, A. & Cotter, J. Lewis Score: a useful clinical tool for patients with suspected Crohn's Disease submitted to capsule endoscopy. *J. Crohns. Colitis* (2012). doi:10.1016/j.crohns.2011.12.002
120. Niv, Y. *et al.* Validation of the Capsule Endoscopy Crohns Disease Activity Index (CECDAI or Niv score): A multicenter prospective study. *Endoscopy* (2012). doi:10.1055/s-0031-1291385
121. Bourreille, A. *et al.* Wireless capsule endoscopy versus ileocolonoscopy for the diagnosis of postoperative recurrence of Crohn's disease: A prospective study. *Gut* (2006). doi:10.1136/gut.2005.081851
122. Pons Beltrán, V. *et al.* Evaluation of postsurgical recurrence in Crohn's disease: a new indication for capsule endoscopy? *Gastrointest. Endosc.* (2007). doi:10.1016/j.gie.2006.12.059
123. Leighton, J. A. *et al.* Comparing diagnostic yield of a novel pan-enteric video capsule endoscope with ileocolonoscopy in patients with active Crohn's disease: a feasibility study. *Gastrointest. Endosc.* (2017). doi:10.1016/j.gie.2016.09.009

124. Wright, E. K. *et al.* Measurement of fecal calprotectin improves monitoring and detection of recurrence of Crohn's disease after surgery. *Gastroenterology* (2015). doi:10.1053/j.gastro.2015.01.026
125. Yamamoto, T., Shimoyama, T., Umegae, S. & Matsumoto, K. Serial monitoring of faecal calprotectin for the assessment of endoscopic recurrence in asymptomatic patients after ileocolonic resection for Crohn's disease: A long-term prospective study. *Therap. Adv. Gastroenterol.* (2016). doi:10.1177/1756283X16646562
126. Boschetti, G. *et al.* Levels of fecal calprotectin are associated with the severity of postoperative endoscopic recurrence in asymptomatic patients with Crohn's disease. *Am. J. Gastroenterol.* (2015). doi:10.1038/ajg.2015.30
127. Garcia-Planella, E. *et al.* Fecal calprotectin levels are closely correlated with the absence of relevant mucosal lesions in postoperative Crohn's Disease. *Inflamm. Bowel Dis.* (2016). doi:10.1097/MIB.0000000000000960
128. Sailer, J. & Vogelsang, H. Anastomotic recurrence of Crohn ' s disease after ileocolic resection : comparison of MR enteroclysis with endoscopy. 2512–2521 (2008). doi:10.1007/s00330-008-1034-6
129. Peyrin-Biroulet, L. *et al.* Azathioprine and 6-Mercaptopurine for the Prevention of Postoperative Recurrence in Crohn's Disease: A Meta-Analysis. *Am. J. Gastroenterol.* (2009). doi:10.1038/ajg.2009.301
130. Regueiro, M. *et al.* In fl iximab Reduces Endoscopic, but Not Clinical, Recurrence of Crohn ' s Disease After Ileocolonic Resection. 1568–1578 (2016). doi:10.1053/j.gastro.2016.02.072
131. Savarino, E. *et al.* Adalimumab is more effective than azathioprine and mesalamine at preventing postoperative recurrence of Crohn's disease: A randomized controlled trial. *Am. J. Gastroenterol.* **108**, 1731–1742 (2013).
132. De Cruz, P. *et al.* Efficacy of thiopurines and adalimumab in preventing Crohn's disease recurrence in high-risk patients - A POCER study analysis. *Aliment. Pharmacol. Ther.* **42**, 867–879 (2015).
133. Vera-Mendoza, I. *et al.* Adalimumab vs azathioprine in the prevention of postoperative Crohn's disease recurrence. A GETECCU randomised trial. *J. Crohn's Colitis* **11**, 1293–1301 (2017).
134. Carla-Moreau, A., Paul, S., Roblin, X., Genin, C. & Peyrin-Biroulet, L. Prevention and treatment of postoperative Crohn's disease recurrence with anti-TNF therapy: A meta-analysis of controlled trials. *Dig. Liver Dis.* (2015). doi:10.1016/j.dld.2014.11.005
135. Singh, S. *et al.* Comparative efficacy of pharmacologic interventions in preventing relapse of Crohn's disease after surgery: A systematic review and network meta-analysis. *Gastroenterology* (2015). doi:10.1053/j.gastro.2014.09.031

136. De Cruz, P. *et al.* Crohn's disease management after intestinal resection: A randomised trial. *Lancet* **385**, 1406–1417 (2015).
137. Yamamoto, T., Umegae, S. & Matsumoto, K. Impact of infliximab therapy after early endoscopic recurrence following ileocolonic resection of Crohn's disease: A prospective pilot study. *Inflamm. Bowel Dis.* (2009). doi:10.1002/ibd.20915
138. Ferrante, M. *et al.* Systematic versus endoscopy-driven treatment with azathioprine to prevent postoperative ileal Crohn's disease recurrence. *J. Crohn's Colitis* (2015). doi:10.1093/ecco-jcc/jjv076
139. Domènech, E. *et al.* Recomendaciones del Grupo Español de Trabajo en Enfermedad de Crohn y Colitis Ulcerosa (GETECCU) sobre la monitorización, prevención y tratamiento de la recurrencia posquirúrgica en la enfermedad de Crohn. *Gastroenterología y Hepatología* **40**, 472–483 (2017).
140. Regueiro, M., Strong, S. A., Ferrari, L. & Fichera, A. Postoperative Medical Management of Crohn's Disease : Prevention and Surveillance Strategies. *J. Gastrointest. Surg.* (2016). doi:10.1007/s11605-016-3172-5
141. Nguyen, G. C., Jr, E. V. L., Hirano, I., Singh, S. & Sultan, S. American Gastroenterological Association Institute Guideline on. 271–275 (2017). doi:10.1053/j.gastro.2016.10.038
142. Mary, J. Y. & Modigliani, R. Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Therapeutiques des Affections Inflammatoires du Tube Digestif (GETAID). *Gut* **30**, 983–989 (1989).
143. Daperno, M. *et al.* Development and validation of a new, simplified endoscopic activity score for Crohn's disease: The SES-CD. *Gastrointest. Endosc.* **60**, 505–512 (2004).
144. Modigliani, R. & Mary JY. Reproducibility of colonoscopic findings in Crohn's disease: a prospective multicenter study of interobserver variation. Groupe d'Etudes Therapeutiques des Affections Inflammatoires du Tube Digestif (GETAID). *Dig. Dis. Sci.* **32**, 1370–9 (1987).
145. Domènech, E. *et al.* Impact of azathioprine on the prevention of postoperative Crohn's disease recurrence: Results of a prospective, observational, long-term follow-up study. *Inflamm. Bowel Dis.* (2008). doi:10.1002/ibd.20359
146. Bayart, P. *et al.* Ileal or Anastomotic Location of Lesions Does Not Impact Rate of Postoperative Recurrence in Crohn's Disease Patients Classified i2 on the Rutgeerts Score. *Dig. Dis. Sci.* **61**, 2986–2992 (2016).
147. Rivière, P. *et al.* No Change in Determining Crohn's Disease Recurrence or Need for Endoscopic or Surgical Intervention With Modification of the Rutgeerts' Scoring System. *Clin. Gastroenterol. Hepatol.* **17**, 1643–1645 (2019).

148. De Cruz, P. *et al.* 312 The First Validated Post-Operative Endoscopic Crohns Disease Index: The POCER Index. Identification of Key Endoscopic Prognostic Factors. *Gastroenterology* (2016). doi:10.1016/s0016-5085(16)30358-4
149. Shah, S. C., Colombel, J. F., Sands, B. E. & Narula, N. Systematic review with meta-analysis: Mucosal healing is associated with improved long-term outcomes in Crohn's disease. *Aliment. Pharmacol. Ther.* **43**, 317–333 (2016).
150. Pineton De Chambrun, G., Peyrin-Biroulet, L., Lémann, M. & Colombel, J. F. Clinical implications of mucosal healing for the management of IBD. *Nature Reviews Gastroenterology and Hepatology* **7**, 15–29 (2010).
151. Reinink, A. R., Lee, T. C. & Higgins, P. D. R. Endoscopic Mucosal Healing Predicts Favorable Clinical Outcomes in Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* **22**, 1859–1869 (2016).
152. De Cruz, P., Kamm, M. A., Prideaux, L., Allen, P. B. & Moore, G. Mucosal healing in Crohn's disease: A systematic review. *Inflamm. Bowel Dis.* **19**, 429–444 (2013).
153. D'Haens, G. R. *et al.* Endpoints for clinical trials evaluating disease modification and structural damage in adults with Crohn's disease. *Inflamm. Bowel Dis.* **15**, 1599–1604 (2009).
154. Sandborn, W. J. *et al.* Ustekinumab induction and maintenance therapy in refractory Crohn's disease. *N. Engl. J. Med.* **367**, 1519–28 (2012).
155. Balzola, F., Cullen, G., Ho, G. T., Russell, R. K. & Wehkamp, J. Adalimumab induces and maintains mucosal healing in patients with Crohn's disease: Data from the EXTEND trial. *Inflamm. Bowel Dis. Monit.* **12**, 151–152 (2012).
156. Ferrante, M. *et al.* Validation of endoscopic activity scores in patients with Crohn's disease based on a post hoc analysis of data from SONIC. *Gastroenterology* **145**, 978–986 (2013).
157. Mary, J. Y., Lemann, M., Colombel, J. F. *et al.* Endoscopic remission and response in Crohn's disease: an objective definition using the CDEIS. *Gut* **54**, A50 (2005).
158. Van Der Woude, C. J. *et al.* Phase I, double-blind, randomized, placebo-controlled, dose-escalation study of NI-0401 (a fully human anti-CD3 monoclonal antibody) in patients with moderate to severe active Crohn's disease. *Inflamm. Bowel Dis.* **16**, 1708–1716 (2010).
159. Hébuterne, X. *et al.* Endoscopic improvement of mucosal lesions in patients with moderate to severe ileocolonic Crohn's disease following treatment with certolizumab pegol. *Gut* **62**, 201–208 (2013).
160. Laharie, D. *et al.* Mucosal healing with methotrexate in Crohns disease: A prospective comparative study with azathioprine and infliximab. *Aliment. Pharmacol. Ther.* **33**, 714–721 (2011).

161. Lémann, M. *et al.* A randomized, double-blind, controlled withdrawal trial in Crohn's disease patients in long-term remission on azathioprine. *Gastroenterology* **128**, 1812–1818 (2005).
162. Hanauer, S. B. *et al.* Maintenance infliximab for Crohn's disease: The ACCENT I randomised trial. *Lancet* **359**, 1541–1549 (2002).
163. Colombel, J. F. *et al.* Infliximab, Azathioprine, or Combination Therapy for Crohn's Disease. *N. Engl. J. Med.* **362**, 1383–1395 (2010).
164. Sandborn, W. J. *et al.* Natalizumab induction and maintenance therapy for Crohn's disease. *N. Engl. J. Med.* **353**, 1912–25 (2005).
165. Louis, E. *et al.* Maintenance of remission among patients with Crohn's disease on antimetabolite therapy after infliximab therapy is stopped. *Gastroenterology* **142**, 63-70.e5 (2012).
166. Colombel, J. *et al.* 2008 IBD Abstracts. **2**, 2008 (2008).
167. Bednar, F., Schofield, H. K., Collins, M. A., Yan, W. & Zhang, Y. *cr ip t pt e d cr t Ac ce pt us.* 1–45 (2015).
168. Baert, F. *et al.* Mucosal Healing Predicts Sustained Clinical Remission in Patients With Early-Stage Crohn's Disease. *Gastroenterology* **138**, 463–468 (2010).
169. Daperno, M. *et al.* Results of the 2nd part Scientific Workshop of the ECCO (II): Measures and markers of prediction to achieve, detect, and monitor intestinal healing in Inflammatory Bowel Disease. *J. Crohn's Colitis* **5**, 484–498 (2011).
170. Vuitton, L. *et al.* IOIBD technical review on endoscopic indices for Crohn's disease clinical trials. *Gut* **65**, 1447–1455 (2016).
171. Cottone, M., Mocciano, F. & Orlando, A. Infliximab prevents Crohn's disease recurrence after ileal resection. *Expert Rev. Gastroenterol. Hepatol.* **3**, 231–234 (2009).
172. Sakuraba, A., Annunziata, M. L., Cohen, R. D., Hanauer, S. B. & Rubin, D. T. Mucosal healing is associated with improved long-term outcome of maintenance therapy with natalizumab in Crohn's disease. *Inflamm. Bowel Dis.* **19**, 2577–2583 (2013).
173. Noman, M. *et al.* Vedolizumab induces long term mucosal healing in patients with Crohn's disease and ulcerative colitis. *J. Crohns. Colitis* **111**, 1147–55 (2017).
174. Khanna, R. *et al.* Reliability among central readers in the evaluation of endoscopic findings from patients with Crohn's disease. *Gut* **65**, 1119–1125 (2016).
175. Daperno, M. *et al.* Inter-observer agreement in endoscopic scoring systems: Preliminary report of an ongoing study from the Italian Group for Inflammatory Bowel Disease (IG-IBD). *Dig. Liver Dis.* **46**, 969–973 (2014).

176. Rutgeerts, P. *et al.* Agreement of site and central readings of ileocolonoscopy scores in Crohn's disease: Comparison using data from the EXTEND trial. *Gastrointest. Endosc.* **83**, 188–197 (2016).
177. Sipponen, T., Nuutinen, H., Turunen, U. & Färkkilä, M. Endoscopic evaluation of Crohn's disease activity: Comparison of the CDEIS and the SES-CD. *Inflamm. Bowel Dis.* **16**, 2131–2136 (2010).
178. Khanna, R. *et al.* Responsiveness of Endoscopic Indices of Disease Activity for Crohn's Disease. *Am. J. Gastroenterol.* **112**, 1584–1592 (2017).
179. Geerse, K. *et al.* Sa1198 Agreement Among Experts in the Endoscopic Evaluation of Postoperative Recurrence in Crohn's Disease Using the Rutgeerts Score. *Gastroenterology* **146**, S-227 (2014).
180. Kennedy, N. A., Ennis, H., Gaya, D. R., Mowat, C. & Arnott, I. D. R. P308. Interobserver agreement in assessment of Rutgeerts' score of endoscopic recurrence of ileal Crohn's disease: a substudy of the TOPPIC trial. *J. Crohn's Colitis* (2015). doi:10.1093/ecco-jcc/jju027.426
181. Marteau, P. *et al.* Interobserver variation study of the rutgeerts score to assess endoscopic recurrence after surgery for Crohn's disease. *J. Crohn's Colitis* **10**, 1001–1005 (2016).
182. Fernandes, S. R., Pinto, J. S. L. D., Marques da Costa, P. & Correia, L. Disagreement Among Gastroenterologists Using the Mayo and Rutgeerts Endoscopic Scores. *Inflamm. Bowel Dis.* **24**, 254–260 (2018).
183. Khanna, R. *et al.* Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review). *Cochrane Database Syst. Rev.* (2016). doi:10.1002/14651858.CD010642.pub2. www.cochranelibrary.com
184. Cder, F. D. A. Clinical Trial Imaging Endpoint Process Standards Guidance for Industry Clinical Trial Imaging Endpoint Process Standards Guidance for Industry. (2018).
185. Administration, D. Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. *Biotechnol. Law Rep.* **26**, 375–386 (2007).
186. Feagan, B. G. *et al.* The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* **145**, 149–157.e2 (2013).
187. Fulton, J. F. History of medical education. *Br. Med. J.* **2**, 457–461 (1953).
188. Rex, D. K. Colonoscopic withdrawal technique is associated with adenoma miss rates. *Gastrointest. Endosc.* **51**, 33–36 (2000).
189. Baxter, N. N. *et al.* Association of colonoscopy and death from colorectal cancer. *Ann. Intern. Med.* **150**, 1–8 (2009).

190. Chen, S. C. & Rex, D. K. Endoscopist can be more powerful than age and male gender in predicting adenoma detection at colonoscopy. *Am. J. Gastroenterol.* **102**, 856–861 (2007).
191. Wigton, R. S., Blank, L. L., Nicolas, J. A. & Tape, T. G. Procedural skills training in internal medicine residencies. A survey of program directors. *Ann. Intern. Med.* **111**, 932–938 (1989).
192. Wigton, R. S., Blank, L. L., Monsour, H. & Nicolas, J. A. Procedural skills of practicing gastroenterologists: A national survey of 700 members of the American College of Physicians. *Ann. Intern. Med.* **113**, 540–546 (1990).
193. Stanford, S. B. Achieving competence in colonoscopy: Milestones and the need for a new endoscopic curriculum in gastroenterology training. *World J. Gastrointest. Endosc.* **7**, 1279 (2015).
194. Walsh, C. M. In-training gastrointestinal endoscopy competency assessment tools: Types of tools, validation and impact. *Best Pract. Res. Clin. Gastroenterol.* **30**, 357–374 (2016).
195. Dubé, C. & Rostom, A. Acquiring and maintaining competency in gastrointestinal endoscopy. *Best Pract. Res. Clin. Gastroenterol.* **30**, 339–347 (2016).
196. Epstein, R. M. Assessment in medical education. *N. Engl. J. Med.* **356**, 387–96 (2007).
197. Forbes, N., Mohamed, R. & Raman, M. Learning curve for endoscopy training: Is it all about numbers? *Best Pract. Res. Clin. Gastroenterol.* **30**, 349–356 (2016).
198. Walsh, C. M., Sherlock, M. E., Ling, S. C. & Carnahan, H. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. in *Cochrane Database of Systematic Reviews* (2012). doi:10.1002/14651858.CD008237.pub2
199. Singh, S., Sedlack, R. E. & Cook, D. a. Effects of Simulation-Based Training in Gastrointestinal Endoscopy: A Systematic Review and Meta-analysis. *Clin. Gastroenterol. Hepatol.* **12**, 1611-1623.e4 (2014).
200. Rex, D. K. *et al.* Quality indicators for colonoscopy. *Am. J. Gastroenterol.* **110**, 72–90 (2015).
201. Joint Advisory Group on GI Endoscopy (UK). JAG Trainee Certification Process - Colonoscopy. **44**, 20–21 (2011).
202. Barton, J. R., Corbett, S. & Van Der Vleuten, C. P. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: Assessing colonoscopic skills of senior endoscopists. *Gastrointest. Endosc.* **75**, 591–597 (2012).
203. Gupta, S. *et al.* Development and validation of a novel method for assessing competency in polypectomy: Direct observation of polypectomy skills. *Gastrointest. Endosc.* **73**, (2011).

204. Sedlack, R. E. The Mayo Colonoscopy Skills Assessment Tool: Validation of a unique instrument to assess colonoscopy skills in trainees. *Gastrointest. Endosc.* **72**, (2010).
205. Sedlack, R. E. Erratum: Training to competency in colonoscopy: Assessing and defining competency standards (Gastrointestinal Endoscopy (2011) 74 (355-366)). *Gastrointestinal Endoscopy* **74**, 729 (2011).
206. Korman, L. Y. *et al.* Capsule Endoscopy Structured Terminology (CEST): Proposal of a standardized and structured terminology for reporting capsule endoscopy procedures. *Endoscopy* **37**, 951–959 (2005).
207. Urien, I. F. Curso cápsula endoscópica Pamplona. Available at: [www.capsulecourse.com/](http://www.capsulecourse.com/).
208. No Title. Available at: <http://www.worldendo.org/education/video-capsule-endoscopy-training-vce>.
209. Courses, U. No Title. Available at: <https://www.ueg.eu/education/online-courses/a-primer-in-capsule-endoscopy/?type=Nv63Afc>.
210. GESA. Available at: <https://www.gesa.org.au/education/capsule-endoscopy-workshop/>.
211. Koffas, A., Laskaratos, F.-M. & Epstein, O. Training in video capsule endoscopy: Current status and unmet needs. *World J. Gastrointest. Endosc.* **11**, 395–402 (2019).
212. Rajan, E. *et al.* Training in small-bowel capsule endoscopy: Assessing and defining competency. *Gastrointest. Endosc.* **78**, 617–622 (2013).
213. Albert, J. G. *et al.* A simple evaluation tool (ET-CET) indicates increase of diagnostic skills from Small bowel capsule endoscopy training courses: A prospective observational european multicenter study. *Med. (United States)* **94**, 1–8 (2015).
214. Fernandez-Urien, I. *et al.* International core curriculum for capsule endoscopy training courses. *Endosc. Int. Open* **05**, E526–E538 (2017).
215. Rubin, D. T. & Hospital, W. Advanced Inflammatory Bowel Disease Fellowships. *Am. J. Gastroenterol.* **108**, 1670–1671 (2013).
216. Daperno, M. *et al.* Training Programs on Endoscopic Scoring Systems for Inflammatory Bowel Disease Lead to a Significant Increase in Interobserver Agreement Among Community Gastroenterologists. *J. Crohn's Colitis* **10**, 1181–1187 (2016). doi:10.1093/ecco-jcc/jjw181
217. Dubcenco, E. *et al.* Effect of standardised scoring conventions on inter-rater reliability in the endoscopic evaluation of Crohn's disease. *J. Crohn's Colitis* **10**, 1006–1014 (2016).



218. Landis, J. R. & Koch, G. G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* (1977). doi:10.2307/2529310
219. McGraw, K. O. & Wong, S. P. Forming Inferences about Some Intraclass Correlation Coefficients. *Psychol. Methods* (1996). doi:10.1037/1082-989X.1.1.30
220. Ma, C. *et al.* Reliability of Endoscopic Evaluation of Postoperative Recurrent Crohn's Disease. *Clin. Gastroenterol. Hepatol.* 4–8 (2019). doi:10.1016/j.cgh.2019.08.046
221. Panés, J., Feagan, B. G., Hussain, F., Levesque, B. G. & Travis, S. P. Central Endoscopy Reading in Inflammatory Bowel Diseases. *J. Crohns. Colitis* **10**, S542–S547 (2016).
222. Gottlieb, K. & Hussain, F. Voting for Image Scoring and Assessment (VISA) - theory and application of a 2 + 1 reader algorithm to improve accuracy of imaging endpoints in clinical trials. *BMC Med. Imaging* **15**, 1–10 (2015).
223. Ahmad, H., Berzin, T. M., Yu, H. J., Huang, C. S. & Mishkin, D. S. Review Central endoscopy reads in inflammatory bowel disease clinical trials : The role of the imaging core lab. **2**, 201–206 (2014).
224. No Title. Available at:  
[http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500003580.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003580.pdf).
225. Waschke, K. A., Anderson, J., Macintosh, D. & Valori, R. M. Training the gastrointestinal endoscopy trainer. *Best Pract. Res. Clin. Gastroenterol.* **30**, 409–419 (2016).





