



# Protocolo de acceso abierto y clasificación de la literatura científica sobre espirometría

## Open access protocol and classification of scientific literature on spirometry

Mario Arturo Flores-Valadez,\* Layla Michán-Aguirre,† Israel Muñoz-Velasco,† Minerva Romero-Pérez†

\*Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas. Ciudad de México, México;

†Universidad Nacional Autónoma de México. Ciudad de México, México.

**RESUMEN.** El personal de salud que realiza pruebas de función respiratoria debe estar informado sobre la literatura especializada para utilizar el equipo, interpretar los resultados y establecer un diagnóstico para el seguimiento de las enfermedades pulmonares. Dentro de estas pruebas, la espirometría es la que tiene una mayor difusión, por lo que es indispensable la capacitación del personal de salud para su correcta realización. El objetivo de esta investigación es crear un protocolo para organizar la literatura técnica y de investigación sobre espirometría, permitiendo un procesamiento rápido y eficiente, tanto para humanos como para máquinas. Se realizaron anotaciones semánticas en 96 documentos especializados sobre espirometría, con 99 etiquetas categorizadas en siete variables clave analizadas: tipo de documento, acceso, tema, pruebas asociadas a la espirometría, etapa implicada en la realización de la espirometría, patrones funcionales y enfermedades estudiadas. Estas anotaciones están disponibles en línea, son de acceso abierto, semánticas e interoperables, y pueden ser procesadas tanto por humanos, como por computadoras en una plataforma amigable (<https://web.hypothes.is/>). Por las características de las anotaciones, médicos y técnicos que realizan espirometrías pueden interactuar con ellas y otros usuarios, con lo que se fomenta el análisis de información clave en salud de manera abierta y social, lo que es de utilidad para la práctica, investigación y enseñanza de la neumología.

**Palabras clave:** datos, hypothes.is, web semántica, espirometría.

**ABSTRACT.** Health personnel who perform lung function tests must be well informed about specialized literature to use the equipment, interpret the results, and establish a diagnosis for monitoring lung diseases. Spirometry is the most widely used lung function test, which is why it is essential to train technical and health personnel for its correct performance. The objective of this research is to create a protocol to organize and classify the technical and research literature on spirometry, allowing fast and efficient processing for both humans and machines. To achieve our objective, semantic annotations were made in 96 specialized documents on spirometry, with 99 tags categorized into seven key variables analyzed: type of document, access, topic, tests associated with spirometry, stage involved in performing spirometry, functional patterns and diseases studied through. These annotations are available online, are open access, semantic and interoperable, and can be processed by both humans and computers on a user-friendly platform (<https://web.hypothes.is/>). Due to the characteristics of the annotations, physicians and technicians who perform spirometry can interact with them and other users, thus promoting the analysis of key health information in an open and social manner, which can be useful for practice, research and teaching of pulmonology.

**Keywords:** Data, hypothes.is, semantic web, spirometry.

### Abreviaturas:

DLCO = difusión pulmonar de monóxido de carbono

PFR = pruebas de función respiratoria

### INTRODUCCIÓN

Las pruebas de función respiratoria (PFR) son fundamentales en el diagnóstico y el seguimiento de las enfermedades

#### Correspondencia:

**Dr. Mario Arturo Flores Valadez**

Departamento de Fisiología Respiratoria,

Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas. Ciudad de México, México.

**Correo electrónico:** [mario.flores.valadez@gmail.com](mailto:mario.flores.valadez@gmail.com)

Recibido: 06-VI-2024; aceptado: 10-II-2025.

**Citar como:** Flores-Valadez MA, Michán-Aguirre L, Muñoz-Velasco I, Romero-Pérez M. Protocolo de acceso abierto y clasificación de la literatura científica sobre espirometría. Neumol Cir Torax. 2024; 83 (3):213-223. <https://dx.doi.org/10.35366/119446>

Artículo open access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)



que afectan los pulmones, ya que nos permiten conocer las propiedades mecánicas, ventilatorias y de intercambio gaseoso, así como otros marcadores de inflamación que se relacionan con el control de las enfermedades.

La espirometría es una de las pruebas con mayor alcance, con una historia de más de 175 años.<sup>1</sup> Esta herramienta ha sido empleada para evaluar el estado de salud y la calidad de vida de los individuos a lo largo del tiempo. A pesar de los avances en la tecnología y en las herramientas, la automatización de estas pruebas es imposible. Por lo tanto, es indispensable la capacitación del personal técnico y en salud para su correcta realización. Aunque los estándares técnicos y de interpretación se actualizan con frecuencia y están disponibles públicamente,<sup>2,3</sup> existen inconvenientes en cuanto a la capacitación y actualización de la información entre el personal que la realiza, requiriendo programas de capacitación continua para lograr un impacto en la aplicación de la prueba,<sup>4</sup> por lo que explorar nuevas formas de transmitir el conocimiento relacionado con esta prueba podría ser una opción para fomentar la difusión de la espirometría.

Convencionalmente, el término espirometría se utiliza para designar a un tipo de PFR en donde se realiza una maniobra forzada. Adicionalmente, existe el término espirometría lenta (*slow vital capacity*, en inglés), en ambas se llenan por completo los pulmones desde capacidad pulmonar total (*total lung capacity*, TLC en inglés), pero en esta última la espiración hasta vaciar los pulmones es a menor velocidad. Para fines de nuestro estudio designamos el término espirometría como sinónimo de espirometría forzada.

El reporte de una espirometría debe contener los siguientes elementos: 1) datos antropométricos y de identificación del paciente; 2) ecuación utilizada para la interpretación; 3) valores obtenidos en las maniobras, la mejor maniobra y su relación con la ecuación de predicción (con el límite inferior de la normalidad [LIN], z-score o porcentaje del predicho); 4) gráficos de las maniobras realizadas; 5) calidad de la maniobra; y 6) en caso de realizar una maniobra posterior a la administración de un medicamento broncodilatador, los mismos datos mencionados para la maniobra posbroncodilatador.<sup>5</sup>

La investigación digital en salud consiste en recopilar, organizar, anotar y analizar información digital, principalmente datos y literatura especializada. En este proceso, se utilizan herramientas informáticas que permiten realizar funciones específicas, generar interoperabilidad y automatizar procesos. Entre las herramientas más comunes se encuentran las bases de datos diseñadas para estructurar y recuperar información;<sup>6</sup> los gestores bibliográficos, que facilitan la tarea de administrar referencias y citas;<sup>7</sup> las alertas y los *feeds*, que permiten enfrentar la inmediatez y la sobrecarga de información mediante la recepción de noticias sobre actualizaciones;<sup>8</sup> y las herramientas

de minería de textos,<sup>9</sup> grandes datos, análisis de redes,<sup>10</sup> cienciometría e inteligencia artificial, las cuales potencian el análisis de la información, tanto en lenguaje natural como en lenguaje máquina.<sup>11</sup>

Además, existen herramientas que funcionan como marcadores y anotadores, permitiendo guardar las direcciones de las páginas web y realizar apuntes en tiempo real en ellas. De tal manera que es posible poner notas en todo tipo de páginas web con texto, artículos, libros, imágenes, videos, audios y conjuntos de datos que se encuentren en la red.

Cuando estas anotaciones incorporan tecnología de la web semántica, se denominan anotadores semánticos. La web semántica es una tecnología promovida por el Consorcio *World Wide Web* (W3C),<sup>12</sup> que permite integrar datos de diferentes tipos, mediante el uso de un formato estandarizado e interoperable a través de enlaces que se aplican en la búsqueda de información, el etiquetado y el diseño de mapas de conocimiento.<sup>13</sup> La anotación facilita la creación, el intercambio y el análisis de los datos generados a partir de los contenidos de las páginas, permitiendo su análisis posterior, tanto cuantitativa como cualitativamente.<sup>14</sup>

La anotación semántica con aplicaciones web abiertas como *hypothes.is* se ha implementado para proyectos de interés en biociencias como el de la revista *e-life*,<sup>15</sup> *scibot*<sup>16</sup> y *sciscore*.<sup>17</sup> Asimismo, se ha utilizado en la investigación clínica<sup>18</sup> y patológica,<sup>19</sup> con resultados prometedores.

Comúnmente, estos temas de innovación tecnológica son técnicos y complejos, y se publica poco sobre ellos en español. Nosotros estamos interesados en difundirlos y presentarlos mediante casos de estudio atractivos e interesantes, aprovechando tecnologías accesibles para todos y herramientas de código abierto.

El objetivo de esta investigación es generar un protocolo para manejar información sobre espirometría en relación con cuatro dimensiones: el equipo, el procedimiento, los resultados y la enfermedad. A través de anotaciones semánticas, buscamos reunir y analizar información clave en salud de manera abierta, social, semántica e interoperable, procesable tanto por humanos como por computadoras. Este enfoque será valioso para la investigación y la enseñanza en nuestra región.

## MATERIAL Y MÉTODOS

El procedimiento de manejo de literatura utilizado en esta investigación sobre PFR (técnica, interpretación y utilidad clínica) se dividió en tres etapas. La primera etapa consistió en la recuperación de literatura relevante. La segunda etapa comprendió la anotación y curación de la literatura científico-técnica. Finalmente, la tercera etapa



Figura 1:

Ejemplo de una anotación generada en Hypothes.is para un documento sobre espirometría.

incluyó el procesamiento, análisis y visualización de los datos obtenidos.

Para localizar la literatura científico-técnica sobre espirometría, consultamos las páginas oficiales de organismos que certifican la calidad de las PFR, como la *European Respiratory Society* (ERS) y la *American Thoracic Society* (ATS). Además, buscamos manuales de equipos frecuentemente comercializados en América Latina.<sup>20,21</sup>

Se seleccionaron documentos clave como ejemplos precisos y pertinentes de la bibliografía, los cuales sirvieron como guía para identificar otros documentos similares y relacionados, un ejemplo es el siguiente documento:<sup>2,3</sup> [https://hyp.is/go?url=https%3A%2F%2Ferj.ersjournals.com%2Fcontent%2F60%2F1%2F2F2101499&group=\\_\\_world](https://hyp.is/go?url=https%3A%2F%2Ferj.ersjournals.com%2Fcontent%2F60%2F1%2F2F2101499&group=__world)

Posteriormente, se identificaron los términos clave utilizados para diseñar filtros y consultas específicas que facilitarían la búsqueda de literatura en las bases de datos académicas. Entre estos términos se incluyeron: «lung function test»[MeSH], «interpretation»[ti], «standard», «technical standard», «manual», «reference values». Se utilizaron cinco plataformas para llevar a cabo las búsquedas: *Google Scholar*, *Europe PMC*, *Scielo*, *Lens* y *PubMed*.

El criterio de inclusión se basó en la recopilación de documentos científico-técnicos que abordan explícitamente la espirometría forzada en humanos,<sup>22</sup> los cuales pueden estar acompañados de otras PFR.<sup>3</sup> Se establecieron cinco criterios de exclusión para los documentos: 1) aquellos que trataban sobre PFR en modelos animales;<sup>23</sup> 2) los

que exploraban PFR experimentales no estandarizadas o realizadas en ambientes poco usuales;<sup>24</sup> 3) documentos centrados en modelos matemáticos y cálculos de ecuaciones;<sup>25</sup> 4) documentos que carecían de variables necesarias para el análisis (consultar más adelante); y 5) aquellos documentos que no abordaban de manera explícita las PFR, sino que se enfocan únicamente en temas relacionados (como principios de fisiología médica, enfermedades respiratorias, salud respiratoria, etcétera), sin abordar aspectos técnicos, teóricos, metodológicos o analíticos de las pruebas.<sup>26-28</sup>

### Anotación y curación

La segunda fase implicó la exploración y anotación de la literatura utilizando Hypothes.is, una plataforma pública y disponible en el enlace: <https://hypothes.is/search?q=tag%3AAnotacionPFR+TipoDePrueba%2Fmesh%2FD002000%2FForcedSpirometry> de modo que todas las anotaciones de los documentos se realizaron en Hypothes.is,<sup>29</sup> un anotador de recursos web que facilita la comprensión de la información, la discusión y la generación de ideas. Es una herramienta de código abierto que permite a los usuarios colocar texto, subrayar, comentar y agregar imágenes o videos sobre la misma página web, emulando la idea de «anotar en los bordes de las páginas del libro», con la ventaja de que las interacciones realizadas por el usuario pueden compartirse en tiempo real, pudiendo ser analizadas por otros usuarios o procesadas por computa-

**Tabla 1:** Siete variables y 99 etiquetas semánticas utilizadas para analizar la literatura científico-técnica sobre espirometría en Hypothes.is.

| Dimensiones   | Variables                             | Anotación  |
|---------------|---------------------------------------|--|
| Equipo        | Tipo de prueba (texto)                | TipoDePrueba/mesh/D002000/ForcedSpirometry<br>TipoDePrueba/mesh/D008451/MaximalVoluntaryVentilation<br>TipoDePrueba/mesh/D000089142/FractionalExhaledNitricOxideTesting<br>TipoDePrueba/mesh/D010993/PlethysmographyWholeBody<br>TipoDePrueba/mesh/D000072277/MaximalRespiratoryPressures<br>TipoDePrueba/mesh/000070857/WalkTest<br>TipoDePrueba/mesh/D011653/PulmonaryDiffusingCapacity<br>TipoDePrueba/mesh/D000403/AirwayResistance<br>TipoDePrueba/mesh/D001985/BronchialProvocationTests<br>TipoDePrueba/mesh/D001784/BloodGasAnalysis   |
|               | Tipo de documento (etiqueta)          | spar/fabio/JournalArticle<br>spar/fabio/ResearchPaper<br>spar/fabio/Book<br>spar/fabio/InstructionManual   |
|               | Tipo de acceso (etiqueta)             | acceso/abierto<br>acceso/cerrado   |
|               | Tema (texto)                          | Tema/mesh/D012137/RespiratorySystem<br>Tema/mesh/D058007/PhysiciansPrimaryCare<br>Tema/mesh/D012890/Sleep<br>Tema/mesh/D010372/Pediatrics<br>Tema/mesh/D013909/Thorax<br>Tema/mesh/D017216/Telemedicine<br>Tema/mesh/D006666/HistoryOfMedicine<br>Tema/mesh/D004389/DurableMedicalEquipment<br>Tema/mesh/D001185/ArtificialIntelligence<br>Tema/mesh/D012016/ReferenceValues<br>Tema/mesh/D000086382/COVID19   |
| Procedimiento | Etapas de la prueba (texto)           | EtapasPrueba/Operativa<br>EtapasPrueba/Estandar<br>EtapasPrueba/Interpretacion<br>EtapasPrueba/Clinica   |
|               | Patrón funcional de la prueba (texto) | PatronFuncional/Obstruccion<br>PatronFuncional/PossibleRestriccion<br>PatronFuncional/PossibleMixto<br>PatronFuncional/Normal<br>PatronFuncional/Broncodilatacion<br>PatronFuncional/NoBroncodilatacion<br>PatronFuncional/RestriccionSimple<br>PatronFuncional/RestriccionCompleja<br>PatronFuncional/TrastornoMixto<br>PatronFuncional/VolumenesNormales<br>PatronFuncional/Hiperinflacion<br>PatronFuncional/PulmonesGrandes<br>PatronFuncional/AumentoFlujoSanguineo<br>PatronFuncional/AnormalidadVascularPulmonar<br>PatronFuncional/PerdidaVolumenLocalizada<br>PatronFuncional/PerdidaAlveoloCapilar<br>PatronFuncional/DifusionNormal<br>PatronFuncional/ImpedanciaNormal<br>PatronFuncional/ObstruccionViaAereaPequeña<br>PatronFuncional/ObstruccionViaAereaTotal<br>PatronFuncional/AlteracionReactancia |

**Continúa la Tabla 1:** Siete variables y 99 etiquetas semánticas utilizadas para analizar la literatura científico-técnica sobre espirometría en Hypothes.is.

| Dimensiones | Variables          | Anotación  |
|-------------|--------------------|--|
|             |                    | PatronFuncional/BajaInflamacionEosinofilica<br>PatronFuncional/ModeradaInflamacionEosinofilica<br>PatronFuncional/AltaInflamacionEosinofilica<br>PatronFuncional/NegativoDCP<br>PatronFuncional/PositivoDCP<br>PatronFuncional/InspiracionMaximaNormal<br>PatronFuncional/InspiraciónMaximaAlterada<br>PatronFuncional/EspiracionMaximaNormal<br>PatronFuncional/EspiracionMaximaAlterada<br>PatronFuncional/RetoBronquialPositivo<br>PatronFuncional/RetoBronquialNegativo<br>PatronFuncional/Normoxemia<br>PatronFuncional/Hipoxemia<br>PatronFuncional/AcidosisRespiratoria<br>PatronFuncional/AcidosisMetabolica<br>PatronFuncional/AlcalosisRespiratoria<br>PatronFuncional/AlcalosisMetabolica   |
| Enfermedad  | Enfermedad (texto) | Enfermedad/mesh/D001249/Asthma<br>Enfermedad/mesh/D001987/Bronchiectasis<br>Enfermedad/mesh/D001991/Bronchitis<br>Enfermedad/mesh/D000086382/COVID19<br>Enfermedad/mesh/D002925/PrimaryCiliaryDyskinesia<br>Enfermedad/mesh/D029424/ChronicObstructivePulmonaryDisease<br>Enfermedad/mesh/D017563/InterstitialLungDisease<br>Enfermedad/mesh/D003550/CysticFibrosis<br>Enfermedad/mesh/D001997/BronchopulmonaryDysplasia<br>Enfermedad/mesh/D009468/NeuromuscularDisease<br>Enfermedad/mesh/D017564/RadiationPneumonitis<br>Enfermedad/mesh/D011009/Pneumoconiosis<br>Enfermedad/mesh/D011649/PulmonaryAlveolarProteinosis<br>Enfermedad/mesh/D013121/SpinalCurvatures<br>Enfermedad/mesh/D012600/Scoliosis<br>Enfermedad/mesh/D054990/IdiopathicPulmonaryFibrosis<br>Enfermedad/mesh/D000081029/PulmonaryArterialHypertension<br>Enfermedad/mesh/D006469/Hemoptysis<br>Enfermedad/mesh/D011655/PulmonaryEmbolism<br>Enfermedad/mesh/D012891/SleepApneaSyndrome<br>Enfermedad/mesh/D008175/LungNeoplasm<br>Enfermedad/mesh/D011656/PulmonaryEmphysema<br>Enfermedad/mesh/D011014/Pneumonia<br>Enfermedad/mesh/D014376/Tuberculosis<br>Enfermedad/mesh/D019896/Alfa1AntitripsinDeficiency<br>Enfermedad/mesh/D011015/PneumoniaAspiration<br>Enfermedad/mesh/D015615/CysticAdenomatoidMalformationLung<br>Enfermedad/mesh/D000092122/BronchiolitisObliteransSyndrome<br>Enfermedad/mesh/D012130/RespiratoryHypersensitivity<br>Enfermedad/mesh/D012829/Silicosis |

doras. Además, la plataforma ofrece la opción de reportar cualquier anotación a los moderadores, quienes revisarán la anotación reportada y determinarán si viola las normas de la comunidad,<sup>29</sup> por lo que fomenta un espacio seguro para la difusión del conocimiento.

### El proceso de anotación

La anotación implica la generación de datos asociados a los artículos para organizarlos y clasificarlos utilizando etiquetas (Figura 1).

Las variables estudiadas abarcaron aspectos relacionados con los documentos y las PFR. En cuanto a los documentos, se consideraron variables como el tipo de documento, el tipo de acceso y el tema. Respecto a las PFR asociadas a la espirometría, se analizaron variables como el tipo de prueba, la etapa implicada en su realización, los patrones funcionales y las enfermedades estudiadas mediante la espirometría. Las variables sobre las PFR se hicieron consistentes con cuatro dimensiones: equipo utilizado para realizar la prueba (espirómetro), procedimiento de realización, resultados desplegados en el reporte y su relación con la enfermedad. Las anotaciones se diseñaron utilizando la etiqueta del proyecto y el nombre de las variables. Para clasificar el tipo de documento, se empleó la ontología *FaBiO*,<sup>30</sup> mientras que, para las variables relacionadas con el tipo de prueba, el tema y la enfermedad, se utilizó el tesauro de *MeSH* (Tabla 1).<sup>31</sup>

### Procesamiento, análisis y visualización

Las anotaciones de Hypothes.is se extrajeron en formato *JSON* y *CSV* mediante dos herramientas digitales diseñadas por Jon Udell: <https://jonudell.info/h/tools.html> y <https://jonudell.info/h/facet>. Posteriormente, los datos fueron procesados mediante programación en *Bash*, se llevaron a cabo los análisis y se realizaron las gráficas correspondientes en R.

Para este propósito, se emplearon los paquetes *readr*, *dplyr*, *ggplot2*, *RColorBrewer* y *Viridis*, los cuales están disponibles para la generación de gráficos y el análisis de resultados.

Además, se creó un repositorio en GitHub <https://github.com/lmichan/PFR> con los detalles de este proyecto y en el que se estarán publicando actualizaciones, datos y el código utilizado. Para las imágenes, se optó por utilizar paletas de colores como *viridis\_d*<sup>32</sup> para garantizar la accesibilidad para personas con discapacidad visual (daltonismo), y la paleta de colores rainbow se reservó para casos en los que se tuvieron múltiples variables.

## RESULTADOS

Se analizaron, recuperaron y anotaron 96 documentos, éstos fueron etiquetados con el nombre del proyecto «AnotacionPFR» y con el texto «TipoDePrueba/mesh/D002000/ForcedSpirometry», en relación con los datos generados de las siete variables y las 99 etiquetas. Estas etiquetas se hicieron de manera consistente con las cuatro dimensiones analizadas (Tabla 1). La colección completa se puede consultar de manera abierta y bajo la licencia *Creative Commons* en el siguiente enlace: <https://hypothes.is/search?q=tag%3AAnotacionPFR+TipoDePrueba%2Fmesh%2FD002000%2FForcedSpirometry>

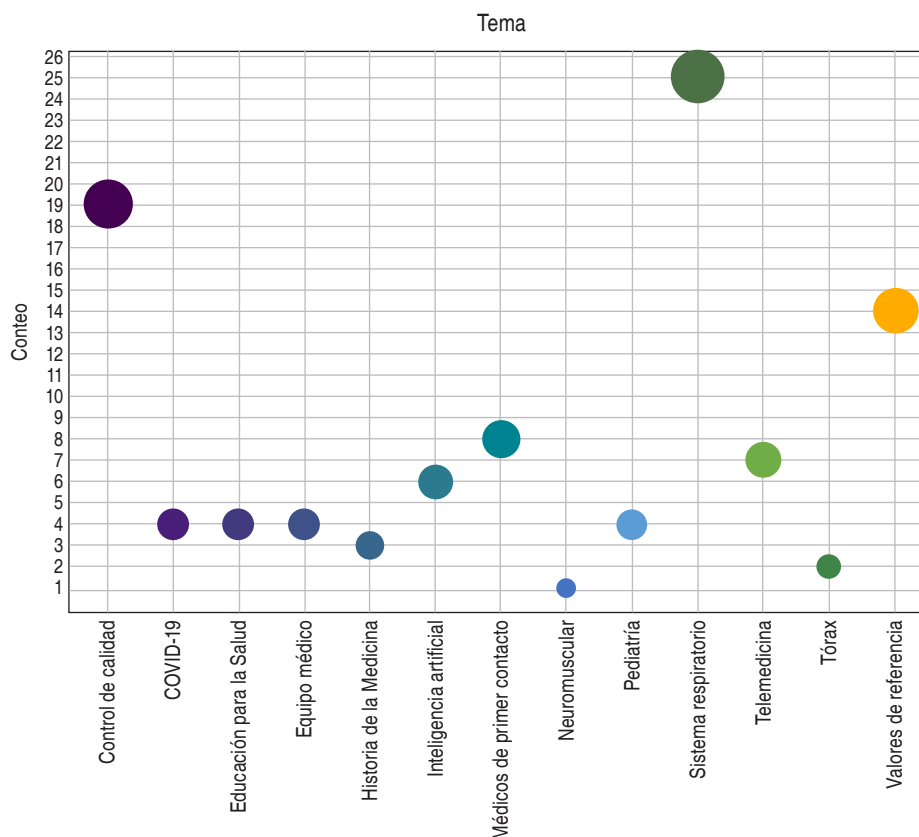
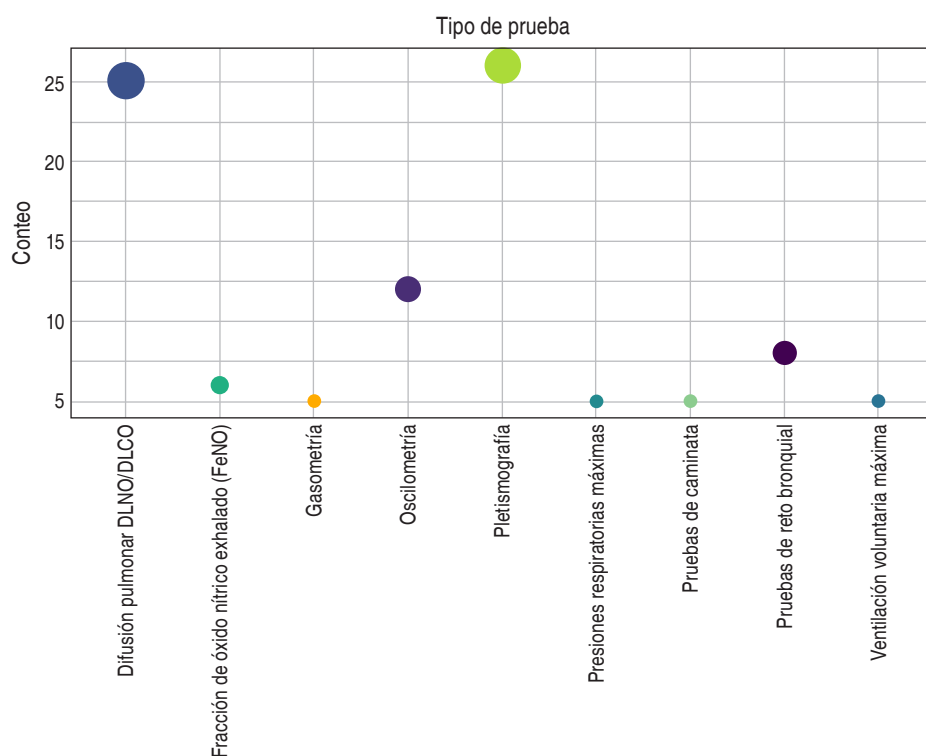


Figura 2:

Temas mencionados en los documentos sobre espirometría.



**Figura 3:**

Tipos de pruebas mencionadas en los documentos sobre espirometría. DLCO = difusión pulmonar de monóxido de carbono (*diffusing capacity of the lungs for carbon monoxide*). DLNO = difusión pulmonar de óxido nítrico (*diffusing capacity of the lungs for nitric oxide*)

De las anotaciones realizadas, se encontró que 78 documentos (81.25%) ofrecían acceso abierto, mientras que 18 documentos (18.75%) requerían una suscripción, 84 fueron artículos de investigación, cuatro libros, dos manuales, dos repositorios de datos, dos sitios web, un catálogo y uno estándar.

Los temas abordados en los documentos fueron 13 en total, esto proporciona información sobre el uso de la espirometría actualmente y la perspectiva de cómo se genera la información (Figura 2). Se observó que el sistema respiratorio (25 artículos), el control de calidad (19 artículos) y los valores de referencia (14 artículos) fueron los tres temas más frecuentes. Los temas restantes como COVID-19, educación para la salud, equipo médico, historia de la medicina, inteligencia artificial, médicos de primer contacto, neuromuscular, pediatría, telemedicina y tórax se distribuyeron entre los otros diez temas analizados.

Los resultados de este análisis incluyen 96 artículos exclusivamente relacionados a la espirometría, entre los cuales también se mencionan otras nueve PFR (Figura 3). En donde la pletismografía, con 26 documentos y la prueba de difusión pulmonar de monóxido de carbono (DLCO), con 25 documentos, son las pruebas más mencionadas; la prueba de reto bronquial que en ocasiones requiere de una maniobra de espirometría forzada se mencionó en ocho documentos. Estos resultados proporcionan una visión general de las pruebas de función respiratoria más comunes y frecuentemente mencionadas en los artículos estudiados.

Se identificaron los patrones funcionales propios de la espirometría en 83 de los documentos, además de que se encontraron otros patrones funcionales de pruebas relacionadas a la espirometría, coincidiendo los patrones de pletismografía y DLCO (Figura 4).

Para estudiar lo que implica hacer una espirometría, clasificamos los documentos según la etapa del estudio, de las cuales se identificaron cuatro: 40 documentos referentes a la interpretación, 19 que mencionan estándares técnicos para realizar la prueba, 19 en de la etapa operativa que se refiere a aspectos técnicos y de equipos, y 14 de la etapa clínica que se relacionaban directamente con el uso de la prueba en la enfermedad.

La última variable evaluada para los documentos fue la relación de la espirometría con distintas enfermedades en la literatura (Figura 5). El asma fue la enfermedad más común, mencionada en 41 documentos, seguida de enfermedad pulmonar obstructiva crónica (EPOC), enfermedad pulmonar intersticial difusa (EPID), enfermedad neuromuscular, fibrosis quística, bronquitis y trastornos de la columna vertebral.

## DISCUSIÓN

En esta investigación creamos un protocolo de anotación para describir y analizar los documentos científicos-técnicos sobre espirometría, esto nos permitió integrar una gran cantidad de información sobre 96 textos relevantes, con 99

etiquetas categorizadas para extraer la información sobre cuatro tipos de documento, dos tipos de acceso, 11 temas de investigación, cuatro etapas implicadas en la realización de la espirometría, 38 patrones funcionales y 30 enfermedades asociadas. Todos estos datos nos permitieron conocer la estructura y las tendencias de esta información, y a partir de la etiqueta empleada para el proyecto (AnotacionPFR), es fácil localizarla de manera simple, precisa y eficiente para que pueda ser reutilizada; con lo que se evita la sobrecarga de información,<sup>33</sup> la infodemia<sup>34</sup> y la desinformación.<sup>35</sup>

Las anotaciones semánticas que se realizan en Hypothes.is permiten aprovechar la tecnología para interactuar con la información, pudiendo extraer los datos publicados y procesarlos por medio de algoritmos para generar nueva información.

Esta plataforma, que puede ser empleada en cualquier área del conocimiento, ha sido utilizada previamente en el área de la salud. En 2016, funcionó para el registro de distintos recursos de investigación (reactivos, materiales y herramientas) utilizadas en los artículos científicos;<sup>36,37</sup> más adelante Goller<sup>38</sup> y colaboradores lo usaron para la enseñanza de la metagenómica Goller;<sup>38</sup> y recientemente Saleipour<sup>39</sup> y su equipo la emplearon para anotar variantes fenotípicas y su correlación genética en una enfermedad rara de amplia heterogeneidad de presentación, la enfermedad de Von Hippel-Lindau (VHL) Saleipour.<sup>39</sup>

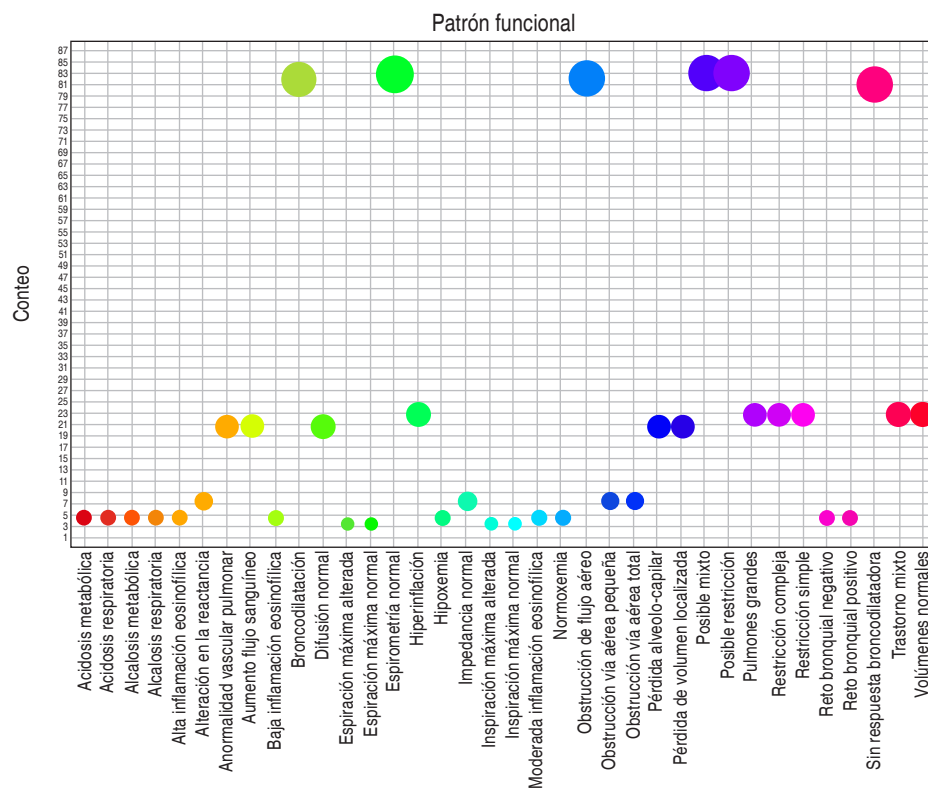
El proceso de diseñar y estructurar las anotaciones fue detallado, largo y minucioso, tuvimos que realizar varias

pruebas para asegurar la congruencia y definir el protocolo correcto. Además, las anotaciones deben ser normalizadas para que la recuperación de los datos generados sea pertinente, consistente y procesada correctamente. Para cumplir con este propósito, utilizamos sistemas de organización del conocimiento (SOC), un tipo de especificación que facilita la modelación de la estructura del significado implícita en un dominio de información, por medio del uso de clases, etiquetas, definiciones, relaciones y propiedades de los conceptos.<sup>40-42</sup> Estas herramientas resultaron sumamente útiles en nuestra investigación para anotar y estructurar datos, facilitaron la búsqueda, la interpretación, el intercambio y la recuperación de contenido digital, además de que nos facilitaron la clasificación de las dimensiones, las categorías y las variables extraídas de los documentos, así como para visualizar el alcance, la estructura, la jerarquía y las relaciones semánticas de las anotaciones, tanto manualmente como por medio de algoritmos. Todo esto está resumido en la [Tabla 1](#) y constituye junto con el protocolo la contribución metodológica más relevante de este artículo.

En la actualidad, el acceso a la información científica y la transparencia en el proceso de generación de conocimiento es muy importante, la accesibilidad y la reproducibilidad de los resultados, tanto de los datos como de los métodos y los resultados de una investigación deben estar disponibles de forma abierta para que otros investigadores puedan acceder a ellos, evaluarlos y replicar los resultados.<sup>43</sup> Herramientas

**Figura 4:**

Patrón funcional mencionado en los documentos sobre espirometría.





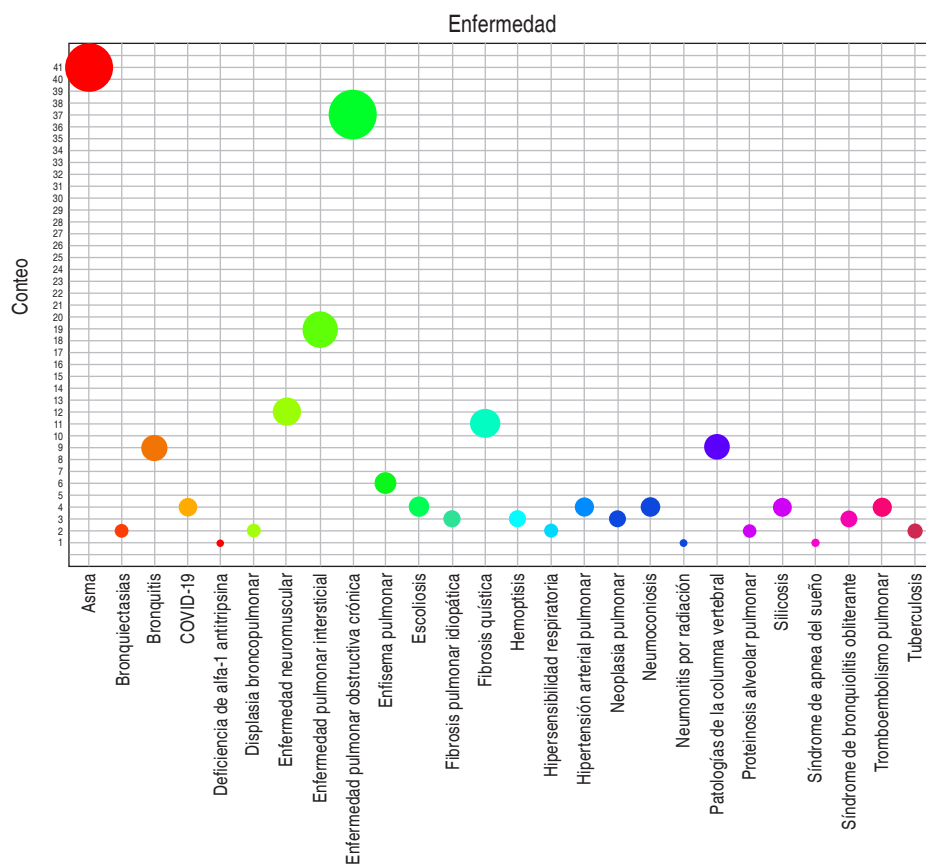


Figura 5:

Tipos de enfermedades mencionadas en los documentos sobre espirometría.

como Hypothes.is fomentan la colaboración y el intercambio de información entre investigadores, lo que promueve un avance más rápido y eficaz del conocimiento científico, los principios de ciencia abierta además permiten que la sociedad en general pueda acceder a la información que sea relevante y confiable.

Como nos interesa practicar y promover la ciencia abierta, nos aseguramos de que la información que generamos estuviera disponible para cualquiera, sin necesidad de crear una cuenta. Para esto se realizaron todas las anotaciones en el grupo público, de tal manera que pueden ser vistas por todos los interesados y son reusables porque tienen una licencia abierta de dominio público.<sup>44</sup> Invitamos a todos los especialistas interesados en espirometría y pruebas de función respiratoria a consultar, reusar y realizar anotaciones para generar un mayor conjunto de información y conocimiento sobre este importante tópico de la neumología.

### Limitaciones

La limitación más importante del protocolo es que el análisis y la selección de los documentos fue realizado únicamente por un médico con alta especialidad en función respiratoria; mientras que la organización y procesamiento de los datos estuvo a cargo de especialistas en el manejo de

la información, esto podría limitar el espectro de la calificación de los artículos y crear un sesgo de selección. Sin embargo, esperamos que, al realizar la invitación para la cooperación con otros especialistas, se pueda generar una mejor retroalimentación de los documentos seleccionados, se puedan compartir ideas, integrar otros documentos al compendio logrado hasta la fecha, con el mismo enfoque para la realización, interpretación y enseñanza de las PFR.

Utilizando la etiqueta principal, al encontrarse en un formato de acceso abierto, cualquiera que tenga acceso a la web de Hypothes.is, se puede agregar información que, como las notas del traductor al pie de página en un libro, aportan más al texto original, con la ventaja de que esta información puede ser procesada también por sistemas informáticos, con la finalidad de promover no sólo la investigación, sino también la difusión y enseñanza de este conocimiento.

### CONCLUSIÓN

Las anotaciones semánticas en páginas como Hypothes.is, permiten la clasificación de los documentos centrados en la espirometría, para que otros usuarios puedan emplearlos con fines educativos y de difusión de esta prueba de función respiratoria.

**Conflicto de intereses:** los autores declaran no tener conflicto de intereses.

## REFERENCIAS

- Kouri A, Dandurand RJ, Usmani OS, Chow CW. Exploring the 175-year history of spirometry and the vital lessons it can teach us today. *Eur Respir Rev.* 2021;30(162):210081. Available from: <https://doi.org/10.1183/16000617.0081-2021>
- Graham BL, Steenbruggen I, Miller MR, Barjaktarevic IZ, Cooper BG, Hall GL, et al. Standardization of spirometry 2019 update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *Am J Respir Crit Care Med.* 2019;200(8):e70-e88. Available from: <https://www.atsjournals.org/doi/10.1164/rccm.201908-1590ST>
- Stanojevic S, Kaminsky DA, Miller MR, Thompson B, Aliverti A, Barjaktarevic I, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J.* 2022;60(1):2101499. Available from: <https://publications.ersnet.org/content/erj/60/1/2101499>
- Benítez-Pérez RE, Vázquez-García JC, Sánchez-Gallén E, Salas-Hernández J, Pérez-Padilla R, et al. Impacto de un programa educativo de espirometría en el primer nivel de atención en México. *Neumol Cir Tórax.* 2021;80(1):29-38. Available from: <https://dx.doi.org/10.35366/99451>
- Culver BH, Graham BL, Coates AL, Wanger J, Berry CE, Clarke PK, et al. ATS Committee on proficiency standards for pulmonary function laboratories. Recommendations for a standardized pulmonary function report. An Official American Thoracic Society technical statement. *Am J Respir Crit Care Med.* 2017;196(11):1463-1472. Available from: <https://www.atsjournals.org/doi/10.1164/rccm.201710-1981ST>
- Collen MF. Specialized Medical Databases. *Comput Med Databases* [Internet]. London: Springer London; 2012. p. 151-182. Available from: [http://link.springer.com/10.1007/978-0-85729-962-8\\_5](http://link.springer.com/10.1007/978-0-85729-962-8_5)
- Lorenzetti DL, Ghali WA. Reference management software for systematic reviews and meta-analyses: an exploration of usage and usability. *BMC Med Res Methodol.* 2013;13(1):141. Available from: <https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-13-141>
- Michán-Aguirre L, Romero-Pérez MM. Inmediatez en salud: la tecnología RSS. *Inv Ed Med.* 2024;13(49):120-128. Available from: <http://riem.facmed.unam.mx/index.php/riem/article/view/1303>
- Luque C, Luna JM, Luque M, Ventura S. An advanced review on text mining in medicine. *WIREs Data Min Knowl Discov.* 2019;9(3):e1302. Available from: <http://riem.facmed.unam.mx/index.php/riem/article/view/1303>
- Kalgotra P, Sharda R. Network analytics: an introduction and illustrative applications in health data science. *J Inf Technol Case Appl Res.* 2023;25(3):305-315. Available from: <https://www.tandfonline.com/doi/full/10.1080/15228053.2023.2187995>
- Basu K, Sinha R, Ong A, Basu T. Artificial intelligence: how is it changing medical sciences and its future? *Indian J Dermatol.* 2020;65(5):365-370. Available from: [https://doi.org/10.4103/ijd.ijd\\_421\\_20](https://doi.org/10.4103/ijd.ijd_421_20)
- World Wide Web Consortium. W3C. W3C. 2024. Available from: <https://www.w3.org/>
- Cheung KH, Prud'hommeaux E, Wang Y, Stephens S. Semantic Web for Health Care and Life Sciences: a review of the state of the art. *Brief Bioinform.* 2009;10(2):111-113. Available from: <https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/bbp015>
- Sakor A, Jozashoori S, Niazmand E, Rivas A, Bougiatiotis K, Aisopos F, et al. Knowledge4COVID-19: a semantic-based approach for constructing a COVID-19 related knowledge graph from various sources and analyzing treatments' toxicities. *J Web Semant.* 2023;75:100760. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1570826822000440>
- Perkel JM. Annotating the scholarly web. *Nature.* 2015;528(7580):153-154. Available from: <https://www.nature.com/articles/528153a>
- RRID Portal. RRID | SciBot. Available from: <https://scicrunch.org/resources/about/scibot>
- Menke J, Roelandse M, Ozyurt B, Martone M, Bandrowski A. The rigor and transparency index quality metric for assessing biological and medical science methods. *iScience.* 2020;23(11):101698. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2589004220308907>
- Mowery DL, Jordan P, Wiebe J, Harkema H, Dowling J, Chapman WW. Semantic annotation of clinical events for generating a problem list. *AMIA Annu Symp Proc.* 2013;2013:1032-1041.
- Wahab N, Miligy IM, Dodd K, Sahota H, Toss M, Lu W, et al. Semantic annotation for computational pathology: multidisciplinary experience and best practice recommendations. *J Pathol Clin Res.* 2022;8(2):116-128. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900128/>
- ndd. EasyOne Air Operator's Manual V1.1. 2018. Available from: [https://henrotech.be/sites/default/files/product/manual/easyone\\_air%20ENG%20Manual.pdf](https://henrotech.be/sites/default/files/product/manual/easyone_air%20ENG%20Manual.pdf)
- ndd. EasyOne ProTM LAB Manual del Operador. 2012. Available from: [https://nddmed.com/\\_Resources/Persistent/6f014bcf8df7622039fb234f96fe70fe6d97667c/easyone-pro-manual-v04b.pdf](https://nddmed.com/_Resources/Persistent/6f014bcf8df7622039fb234f96fe70fe6d97667c/easyone-pro-manual-v04b.pdf)
- García-Río F, Calle M, Burgos F, Casan P, del Campo F, Galdiz JB, et al. Spirometry. *Arch Bronconeumol.* 2013;49(9):388-401. Available from: <http://archbronconeumol.org/en-spirometry-articulo-S1579212913001341>
- Devos FC, Maaske A, Robichaud A, Pollaris L, Seys S, Lopez CA, et al. Forced expiration measurements in mouse models of obstructive and restrictive lung diseases. *Respir Res.* 2017;18(1):123. Available from: <http://respiratory-research.biomedcentral.com/articles/10.1186/s12931-017-0610-1>
- Prisk GK, Fine JM, Cooper TK, West JB. Vital capacity, respiratory muscle strength, and pulmonary gas exchange during long-duration exposure to microgravity. *J Appl Physiol.* 2006;101(2):439-447. Available from: <https://www.physiology.org/doi/10.1152/japplphysiol.01419.2005>
- Guiard Y. Understanding the within-individual variability of forced vital capacity: an exploitation of the nhanes iii spirometry data. 2021. Available from: <https://hal.science/hal-03316189>
- Feher J. Lung volumes and airway resistance. *Quant Hum Physiol.* Elsevier; 2012. p. 563-571. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B978012382163800061X>
- Nichols DP. Functional assessment of asthma. *Pediatr Allergy Princ Pract.* Elsevier; 2016. p. 267-275.e2. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780323298759000306>
- Narayanan M, Silverman M. Long-term consequences of childhood respiratory disease. *Kendig Chernick's Disord Respir Tract Child.* Elsevier; 2012. p. 278-283. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9781437719840000176>
- Hypothesis. Hypothesis. 2023. Available from: <https://web.hypothes.is/>

30. Peroni S, Shotton D. FaBiO and CiTO: ontologies for describing bibliographic resources and citations. *J Web Semant.* 2012;17:33-43. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1570826812000790>
31. National Library of Medicine. Medical Subject Headings. 2024. Available from: <https://www.nlm.nih.gov/mesh/meshhome.html>
32. Garnier S, Ross N, Rudis B, Filipovic-Pierucci A, Galili T, timelyportfolio, et al. sjmgarnier/viridis: CRAN release v0.6.3. Zenodo; 2023. Available from: <https://zenodo.org/record/7890878>
33. Siegel MG, Rossi MJ, Lubowitz JH. Artificial intelligence and machine learning may resolve health care information overload. *Arthrosc J Arthrosc Relat Surg.* 2024;40(6):1721-1723. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0749806324000124>
34. Choi S. The coronavirus disease 2019 infodemic: a concept analysis. *Front Public Health.* 2024;12:1362009. Available from: <https://www.frontiersin.org/articles/10.3389/fpubh.2024.1362009/full>
35. Ishizumi A, Kolis J, Abad N, Prybylski D, Brookmeyer KA, Voegeli C, et al. Beyond misinformation: developing a public health prevention framework for managing information ecosystems. *Lancet Public Health.* 2024;9(6):e397-e406. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2468266724000318>
36. Bandrowski A, Brush M, Grethe JS, Haendel MA, Kennedy DN, Hill S, et al. The resource identification initiative: A cultural shift in publishing. *Neuroinform.* 2016;14(2):169-182. Available from: <http://link.springer.com/10.1007/s12021-015-9284-3>
37. Judell. SciBot: Machine and human annotators working together. Hypothesis. 2016. Available from: <https://web.hypothes.is/blog/introducing-developer-api-tokens/>
38. Goller CC, Vandegrift M, Cross W, Smyth DS. Sharing notes is encouraged: annotating and cocreating with Hypothes.is and Google Docs. *J Microbiol Biol Educ.* 2021;22(1):ev22i1.2135. Available from: <https://journals.asm.org/doi/10.1128/jmbe.v22i1.2135>
39. Salehipour D, Farncombe KM, Andric V, Ansar S, Delong S, Li E, et al. Developing a disease-specific annotation protocol for VHL gene curation using Hypothes.is. *Database.* 2023;2023:baac109. Available from: <https://doi.org/10.1093/database/baac109>
40. Zeng ML. Knowledge Organization Systems (KOS). *Knowl Organ.* 2008;35(2-3):160-182. Available from: <https://www.nomos-elibrary.de/index.php?doi=10.5771/0943-7444-2008-2-3-160>
41. Mazzocchi F. Knowledge Organization System (KOS): an introductory critical account. *Knowl Organ.* 2018;45(1):54-78. Available from: <https://www.nomos-elibrary.de/index.php?doi=10.5771/0943-7444-2018-1-54>
42. Hodge GM. Systems of knowledge organization for digital libraries: beyond traditional authority files. *Digital Library Federation*; 2000.
43. Reichmann S, Wieser B. Open science at the science-policy interface: bringing in the evidence? *Health Res Policy Syst.* 2022;20(1):70. Available from: <https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-022-00867-6>
44. Hypothesis. What is the license on annotations? Hypothesis. 2023. Available from: <https://web.hypothes.is/help/what-is-the-license-on-annotations>