

## Review Article

# Methods for identifying biomedical translation: a systematic review

Javier Padilla-Cabello<sup>1,2</sup>, Antonio Santisteban-Espejo<sup>3,4</sup>, Ruben Heradio<sup>5</sup>, Manuel J Cobo<sup>6</sup>, Miguel A Martin-Piedra<sup>7</sup>, Jose A Moral-Munoz<sup>4,8</sup>

<sup>1</sup>Program of Biomedicine, University of Granada, Granada, Spain; <sup>2</sup>Department of Otorhinolaryngology, Hospital Universitario Clínico San Cecilio, Granada, Spain; <sup>3</sup>Department of Pathology, Puerta del Mar University Hospital, Cadiz, Spain; <sup>4</sup>Institute of Research and Innovation in Biomedical Sciences of The Province of Cadiz (INiBICA), Cadiz, Spain; <sup>5</sup>School of Computer Science, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain; <sup>6</sup>Department of Computer Science and Engineering, University of Cadiz, Puerto Real, Spain; <sup>7</sup>Tissue Engineering Group, Department of Histology, University of Granada, Granada, Spain; <sup>8</sup>Department of Nursing and Physiotherapy, University of Cadiz, Cadiz, Spain

Received July 27, 2021; Accepted December 30, 2021; Epub April 15, 2022; Published April 30, 2022

**Abstract:** Translational medicine is an important area of biomedicine, and has significantly facilitated the development of biomedical research. Despite its relevance, there is no consensus on how to evaluate its progress and impact. A systematic review was carried out to identify all the methods to evaluate translational research. Seven methods were found according to the established criteria to analyze their characteristics, advantages, and limitations. They allow us to perform this type of evaluation in different ways. No relevant advantages were found between them; each one presented its specific limitations that need to be considered. Nevertheless, the Triangle of Biomedicine could be considered the most relevant method, concerning the time since its publication and usefulness. In conclusion, there is still a lack of a gold-standard method for evaluating biomedical translational research.

**Keywords:** Translational research, research evaluation, methods for evaluating, bibliometric analysis, biomedical translational research

### Introduction

Translational medicine is an important area of biomedicine that has significantly facilitated the development of biomedical research [1]. The National Center for Advancing Translational Science (NCATS) considers biomedical translation as “the process of turning observations in the laboratory, clinic and community into interventions that improve the health of individuals and the public—from diagnostics and therapeutics to medical procedures and behavioral changes”. It also defines translational science as “the field of investigation focused on understanding the scientific and operational principles underlying each step of the translational process” [2]. Consequently, this field holds a key place in the present and future development and dissemination of interventions that improve human health [3].

The National Institute of Health (NIH) pays special attention to translational research, consid-

ering it as a focus area, as shown by its substantial investment in the Clinical and Translational Science Award (CTSA) program. The CTSA program is one of the most important initiatives in translational medical funding to date [4], and is producing a strong and growing body of influential research results [5]. In 2010, the UK increased the budget of the Medical Research Council (MRC), intending to enhance support for translational research [6]. Similarly, in 2013, the EU created the European Infrastructure for Translational Medicine (EATRIS), a network of European biomedical translation hubs incorporating over eighty academic research centers [7, 8].

As a result, the optimization of generating science applicable to the clinical setting is gaining popularity and generating interest in the scientific community [9]. However, it is necessary to assess whether this effort achieves its objective. Aiming to promote biomedical translation,

## Identification of biomedical translation

each institution should evaluate its progress and its impact, identifying the different scientific outputs generated by translational research.

In this sense, many authors have provided methods for evaluating translational progress and its outcomes in recent years. Thus, evaluators should consider multiple methods of identifying translational outcomes, and there should be a continuous measurement of all methods and metrics over time [10].

One of the first attempts to measure translation was the Translational Research Impact Scale (TRIS) [11]. The TRIS is a standardized measurement tool that provides a practical and objective evaluation of the eventual impact of translational research. The development of TRIS could be the first attempt to get a uniform criterion that allows comparison of outcomes or effects of translational research between institutions. This kind of tool could solve some problems related to the great heterogeneity that we face in translational science. However, the development of these tools must be made as resistant to gaming as possible, as they may influence the metrics used [12].

Based on the available literature, several sources and measures were used to develop this task. There have been several methodologies reported for evaluating translational research, including process analysis, cost analysis, or research publication analysis, among other methods [13]. There is also a protocol for conducting case studies to study the multifaceted processes of translating research processes underlying the development of successful health interventions [14]. Other studies are focused on specific bibliographical-based approaches. In this milieu, Grant *et al.* used the analysis of the bibliography of clinical guidelines as a bibliometric approach to assess the impact of laboratory research on clinical practice [15]. Also, Williams *et al.* proposed data mining and network analysis to identify and link scientific discoveries with advances in clinical medicine [16]. Luke *et al.* presented the Translational Science Benefits Model (TSBM), a new framework that can measure broader impacts of translational research beyond bibliometric measures and scholarly outcomes, such as lives saved, cost savings, or improvements to health [17]. Last, Rollins *et al.* used a

payback framework (five categories: knowledge, research targeting, capacity building, and absorption, policy and product development and broader economic benefits) and a case study model with multiple complementary bibliometric measures to measure and track the long-term and diverse outcomes of CTSA pilot projects [18].

However, this type of analysis does not properly assess the translational component of research. In this sense, some methods and theoretical approaches have been focused on the categorization and classification of journal articles and, consequently, research areas, along the spectrum of translational research. In this sense, these attempts to categorize article journals in terms of their translational features could also apply to a corpus of documents of the same research area and, thus, help assess its translational profile.

The main aim of this work is to perform a systematic review to identify and gather all the publications that describe translational evaluation methods, exposing the characteristics of each one, and their advantages and limitations. The reported findings should facilitate further research in translational research evaluation.

### Methods

#### *Search strategy*

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines were followed to perform this systematic review [19]. The literature search was performed in December 2020 in the following electronic databases: Web of Science (WoS), Scopus, and Medline (using WoS interface). The search queries employed in each database were detailed in **Table 1**.

#### *Selection criteria*

The following inclusion criteria were established to define which documents and issues were of interest for our systematic review: (i) articles describing a methodology of translational research assessment; and (ii) those methodologies based on bibliographical metadata. Moreover, the following exclusion criteria were applied when full-text articles were reviewed: (i) application of a methodology, pre-

# Identification of biomedical translation

**Table 1.** Search strategies and results in each database

| Database | Search strategy   | Results |
|----------|---|---------|
| WoS      | TI = (“translational research” OR “translational science” OR “translational progress”) AND TS = (bibliomet* OR scientomet* OR “big data” OR “big-data” OR “machine learning” OR “machine-learning” OR “science of science”)               | 38      |
| Scopus   | TITLE (“translational research” OR “translational science” OR “translational progress”) AND TITLE-ABS-KEY (bibliomet* OR scientomet* OR “big data” OR “big-data” OR “machine learning” OR “machine-learning” OR “science of science”)     | 45      |
| Medline  | (“translational research” [Title] OR “translational science” [Title] OR “translational progress” [Title]) AND (bibliomet* OR scientomet* OR “big data” OR “big-data” OR “machine learning” OR “machine-learning” OR “science of science”) | 43      |

viously described, to a particular research field; (ii) those presenting web platforms applying the methodologies described in other articles; and (iii) the translational research appears as an aspect of the topic covered by the article, but it was not measured.

### *Study selection process and data extraction*

The literature search was carried out by combining keywords in the scientific databases mentioned above, and duplicated articles were excluded. Next, titles and abstracts were reviewed, and we excluded those articles that did not meet the established inclusion criteria. The remaining articles were analyzed strictly and were finally included in the systematic review. Four reviewers (JP-C, AS-E, MAM-P, and JAM-M) took part independently in the study selection process, review, and systematic data extraction. This process was performed using the Rayyan system for systematic reviews [20]. All the reviewers carried out the selection process in blinded mode. Subsequently, the conflicting registries were discussed to reach a consensus. Furthermore, to obtain further relevant documents, the bibliography of the documents screened in the process detailed below was also reviewed. The detected documents were considered as “Additional records identified through other sources”. The following data were extracted from the studies: author, classification methods, source, unit of analysis, level of analysis, categorization, method basis, translation reference, and main results.

### **Results**

The search strategy proposed in the Methodology section retrieved 126 possibly relevant documents.

Once metadata were retrieved from all documents, duplicates were discarded, and the remaining 45 articles were submitted to title/abstract screening. 34 documents were also discarded as they did not fulfill inclusion criteria

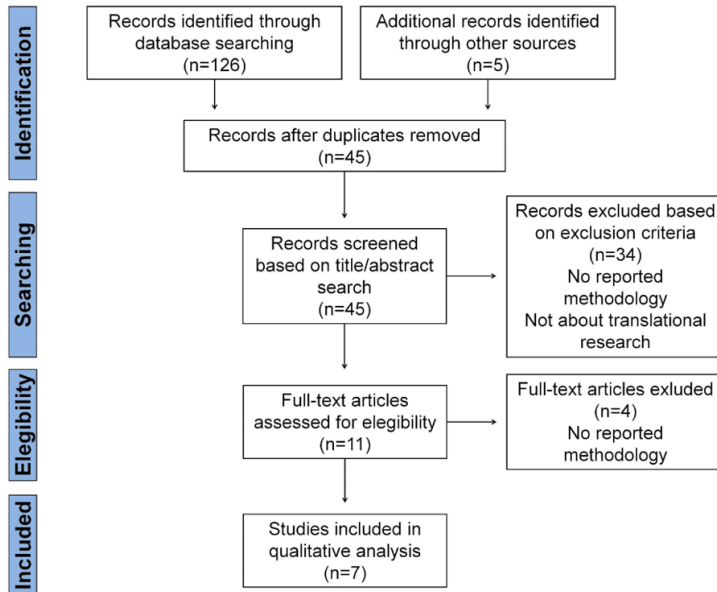
(i.e., they were not directly aimed at assessing the translational profile of research or they were not describing a method for this purpose). After accessing full text, the other four articles were withdrawn since they were not reporting a reproducible method for translational assessment. Finally, seven studies were eligible for critical review and qualitative synthesis (**Figure 1**).

After analysis and extraction of information about the seven studies, two different classifications of the methods could be established, based on the use of descriptors or citations to assess the translational features of the unit of analysis (**Table 2**). Likewise, the descriptors-based methods are subdivided into controlled (e.g., MeSH terms) and uncontrolled descriptors (e.g., Title or abstract). Based on these results, four of the seven reviewed methods reported a translational identification based on the analysis of descriptor terms, two of them using MeSH terms (controlled terms) and the other two using other uncontrolled descriptors. Moreover Three of the methods were reported that attempted to assess translational features of research using citation measurements.

### *Classification methods based on controlled descriptors*

Several methods have been described to establish whether a study is translational in recent years. One of the most cited and relevant methods, and the first to be reported, is the Triangle of Biomedicine (**Figure 2**) [21], which is a graphical classification method comprising an equilateral triangle with three topic areas, one in each of the triangle’s corners: Animals and other complex organisms (A), Cells and molecules (C), and Humans (H). MeSH terms of each article are used to place each publication in this trilinear space. This graphical approach allows grouping the publications and placing them as a whole, for example, as a research area, as well as the assessment of the trends in the groups of publications over

## Identification of biomedical translation



**Figure 1.** Study selection flow diagram for the present systematic review. The literature search on Web of Science, Scopus and Medline databases was performed in December 2020 according to PRISMA guidelines.

time. In this sense, if a research area moves toward the H corner over time, its translational level can be considered higher. One of the most important features of this method, and a clear advantage it has over quantitative approaches (i.e., TS score or APT index), is the graphical representation that gives us relevant information in a clear and easy to interpret way. The position, the size, and the color of each circle give us visual information about the average position of the publications in a discipline (**Figure 2A**), the proportional number of publications in that discipline, and the Translational Distance (the average number of citations generations required to reach an H article), respectively. The Triangle of Biomedicine is also valuable to graphically capture and track changes over time of a topic or research area (**Figure 2B**). Moreover, Weber *et al.* introduced the concept of “generations” of translation lag: an article categorized as A or C that receives a direct quote from an H publication is considered as the “first generation”; if it was cited by another A or C article that was cited by an H publication, it would be the “second generation” and subsequently [21].

This trilinear approach has been the basis of other reports that have attempted to improve its feasibility. In fact, Ke developed a quantitative method that allows the identification of

translational research measuring the “basicness” of an article, a journal, disciplines, or even the entire biomedical literature by learning embeddings of controlled vocabulary [22]. Ke propounded a bottom-up approach: the position of an article on the basic-applied research spectrum is based on the average positions of its MeSH terms. As we can guess, this work was inspired by Weber’s idea that controlled keywords (MeSH) can determine which level of research the article belongs to [21]. Through complex calculations, this results in the Level Score (LS), a continuous variable ranging from -1 (article more orientated toward basic research) to 1 (more applied research). This measurement was consistent with the previously reported methods [22].

### *Classification methods based on uncontrolled descriptors*

Two other classification methods were retrieved in the literature based on uncontrolled descriptors, mainly words from author keywords, title or abstract, that are able to be processed and weighted by machine learning methods to estimate their translational features.

One of these methods, developed by Surkis *et al.*, consists of the characterization of each stage of the translational spectrum (NCTSA T0-T4) through a definitional checklist for each one of them. Articles can be categorized in different translational stages by applying machine learning-based text classifiers. The algorithm used is a Bayesian logistic regression that processes texts from different sections of the articles, such as the title, abstract, and full MeSH terms/phrases, and then classifies them along the translational research spectrum [23].

Similarly, convolutional neural networks (i.e. a type of artificial intelligence neural network) were used to identify possible translational research in the genomics research area, classifying and prioritizing these articles to help curators successfully extract and label the later translational applications of genomics [24]. These methods showed high accuracy in predicting the translational profile of an investiga-

## Identification of biomedical translation

**Table 2.** Main characteristics of the methodologies included in the systematic review

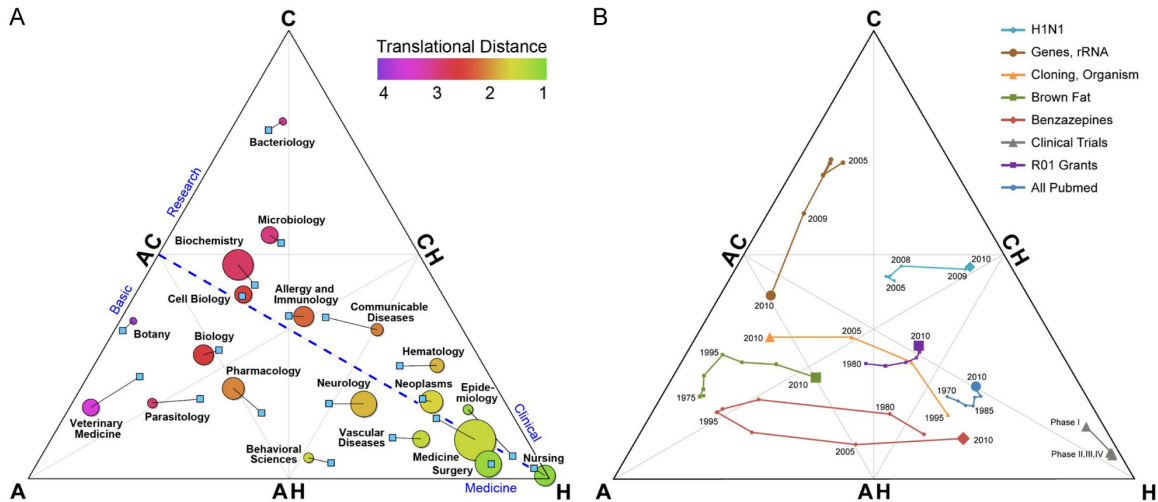
| Author                 | Classification methods     | Source                       | Unit of analysis   | Level of analysis | Categorization   | Method Basis  | Translation reference   | Main results   |
|------------------------|----------------------------|------------------------------|--|-------------------|--|---|---|--|
| Weber (2013)           | Descriptors - controlled   | PubMed                       | MeSH descriptors   | Article           | <ul style="list-style-type: none"> <li>● Humans (H): subtrees B01.050.150.900.649.801.400.112.400.400 (Human) and M01 (Person).</li> <li>● Animals and other complex organisms (A): subtree B01 (Eukaryota) except the code for Humans.</li> <li>● Cells and molecules (C): subtrees A11 (Cells), B02 (Archaea), B03 (Bacteria), B04 (Viruses), G02.111.570 (Molecular Structures), and G02.149 (Chemical Processes).</li> </ul>   | A Cartesian system in which each corner (A, C and H) is a distance of 1 from the origin. The Translational Axis (TA) is a line from the AC point, through the origin, to the H corner. The position of a point projected onto the TA is its Translational Index (TI).   | Documents with MeSH terms included in the H category.   | About 19% of articles could not be classified by this method.<br>Good correlation with Narin's classification schemes*.  |
| Surkis et al. (2016)   | Descriptors - uncontrolled | PubMed                       | Text from title, abstract, and full MeSH terms/phrases associated with a publication | Article           | <ul style="list-style-type: none"> <li>● T0 (basic biomedical research).</li> <li>● T1 (translation to humans).</li> <li>● T2 (translation to patients).</li> <li>● T3 (translation to practice).</li> <li>● T4 (translation to communities).</li> </ul>   | <ol style="list-style-type: none"> <li>1. <u>Manual training</u>: 200 CTSA-granted documents encoded by experts from different CTSA institutions.</li> <li>2. <u>Pre-processing</u>: Text classification weighting words in terms of frequency and location.</li> <li>3. <u>Machine learning algorithms</u>:               <ol style="list-style-type: none"> <li>a. Naïve Bayes (NB).</li> <li>b. Bayesian logistic regression (BLR).</li> <li>c. Random forests (RF).</li> <li>d. Support vector machines (SVM).</li> </ol> </li> </ol> | Documents aimed at health practice to the population, providing communities with the optimal intervention.  | Performed algorithms showed variable usefulness for classification in each category: <ul style="list-style-type: none"> <li>● RF and SVM showed the best accuracy for T0.</li> <li>● RF, BLR and SVM showed the best accuracy for T1/T2.</li> <li>● BLR and SVM showed the best accuracy for T3/T4.</li> </ul> |
| Han et al. (2018)      | Citation                   | Pubmed Scopus                | Publication types Citations  | Article           | <ul style="list-style-type: none"> <li>● Primary translational research</li> <li>● Secondary translational research</li> </ul>   | <ol style="list-style-type: none"> <li>1. <u>Data collection</u>: MeSH publication types and citation metadata were retrieved from PubMed</li> <li>2. <u>Classification</u>:               <ol style="list-style-type: none"> <li>a. Primary translational research. Documents with MeSH publication type corresponding to a clinical study (interventional or observational).</li> <li>b. Secondary translational research. Documents cited by primary translational research.</li> </ol> </li> </ol>                                    | Documents with MeSH publication types corresponding to clinical studies.  | This approach corrects the overestimation of H categorization that sometimes occurs with the Triangle of Biomedicine.  |
| Hutchins et al. (2019) | Citation                   | PubMed iCite (Citation data) | MeSH terms Relative Citation Rate (RCR)  | Article           | <ul style="list-style-type: none"> <li>● H, C and A (as defined by Weber, 2013).</li> <li>● Disease (D)               <ul style="list-style-type: none"> <li>○ Binary: C branch of the MeSH tree, except C22 (Animal Diseases)</li> </ul> </li> <li>● Therapeutic/Diagnostic Approaches (E)               <ul style="list-style-type: none"> <li>○ Binary: E branch of the MeSH tree, except E07 (Equipment and Supplies)</li> </ul> </li> <li>● Chemical/Drug (CD)               <ul style="list-style-type: none"> <li>○ Binary: D branch of the MeSH tree.</li> </ul> </li> </ul> | <ol style="list-style-type: none"> <li>1. <u>Training</u>: Articles/Citing articles stats, RCR and citation by clinical documents in 100,000 documents.</li> <li>2. <u>Machine learning algorithm</u>: Random forest.</li> <li>3. Approximate potential to translate (APT score): prediction of a document to be cited by clinical trials or guidelines binned into 5 groups: &gt;95%, 75%, 50%, 25%, and &lt;5%.</li> </ol>  | A translational document is a document cited by a clinical trial or guideline. There is an increase in the translation potential of an article when it moves from one APT score bin to a higher one in the subsequent year. | The model accurately predicts whether a research paper will eventually be cited by a clinical article after 2 years of publication. MeSH terms and RCR are not eligible by authors (gaming-resistant framework).   |

## Identification of biomedical translation

|                   |                            |  |  |         |   |  |   |   |
|-------------------|----------------------------|--|--|---------|---|--|---|---|
| Hsu et al. (2019) | Descriptors - uncontrolled | Office of Public Health Genomics (OPHG). | Word features from articles (title and abstract) | Article | <ul style="list-style-type: none"> <li>Initial bench-to-bedside phases (T1)</li> <li>Beyond bench-to-bedside phases (T2-T4).</li> </ul>   | <ol style="list-style-type: none"> <li><u>Manual Training</u>: 2286 articles annotated by CDC curators, according to Clyne et al., 2014**.</li> <li>Supervised machine learning algorithm:               <ol style="list-style-type: none"> <li>Convolutional neural networks (CNNs)</li> <li>Support vector machine (SVM)</li> </ol> </li> </ol>  | Documents aimed at the evaluation, utility, and implementation of evidence-based applications into clinical practice and evaluation of the impact these applications have on population health? | High classification accuracy. CNN predicts a better translational profile of an article than SVM. Difficult to distinguish among T2, T3 and T4 levels: low amount of training documents.  |
| Ke (2019)         | Descriptors - controlled   | PubMed WoS (Citation data)               | MeSH terms                                       | Article | Modification of the classification proposed by Weber, 2013) [21]: <ul style="list-style-type: none"> <li>Basic: Cell (C) and animal (A) nodes terms</li> <li>Applied: Human (H) nodes.</li> </ul>   | <ol style="list-style-type: none"> <li>Compute time-evolving co-occurrence matrices <i>between</i> MeSH terms and embedding matrices into d-dimensional vector space.</li> <li>Obtention of Translational Axis (TA), the vector from basic to applied terms.</li> <li><u>Level Score (LS)</u>: is the cosine similarity between each term's vector and the TA vector. LS of a document is the average LS of its MeSH terms at time t.</li> </ol> | LS = +1 (max. value)<br>An article is more or less translational according to the similarity of the MeSH terms assigned to this record compared to basic or applied MeSH.                       | Good correlation with Narin's 4-level journal categorization* and with Weber's analysis. Bimodal distribution (basic/applied) of Medline literature research. (Threshold: LS =0.16). Direct citations rarely occurred between basic and applied research. |
| Kim et al. (2020) | Citation                   | Web of Science                           | Journal's field of knowledge Citations           | Journal | Matching of Web of Science Categories (WoSC) and National Science Foundation (NSF) classification of fields of study <ul style="list-style-type: none"> <li>Clinical research:               <ul style="list-style-type: none"> <li>Clinical science (Clinical).</li> </ul> </li> <li>Non-clinical research:               <ul style="list-style-type: none"> <li>Non-clinical science (Non-Clinical).</li> <li>Multidisciplinary (Multi).</li> <li>Non-science and engineering (Non-S&amp;E).</li> </ul> </li> </ul> | $TS\ score = \frac{Clinical\ citations}{Total\ citations}$ Clinical citations: forward citations from documents published in clinical research journals  | Translational features of non-clinical research are determined by the share of forwarding citations from publications in a clinical research journal.   | TS score changes over time. Possibility to track short/long term impact. High sensitivity: reliable measurement with 5 or more citations. Time lag for reliable measurement (citation-dependent).   |

\*Refers to "Narin F, Pinski G and Gee HH. Structure of the biomedical literature. J Am Soc Inf Sci 1976; 27: 25-45". \*\*Refers to "Clyne M, Schully SD, Dotson WD, Douglas MP, Gwinn M, Kolor K, Wulf A, Bowen MS and Khoury MJ. Horizon scanning for translational genomic research beyond bench to bedside. Genet Med 2014; 16: 535-538".

## Identification of biomedical translation



**Figure 2.** Triangle of Biomedicine. A. Different disciplines located along with the triangle; B. Tracking changes of different topics or research areas over time. Figures obtained from Weber, 2013 [21].

tion. However, these machine-learning-based methods employed supervised learning, requiring a training set of translational documents. Thus, the accuracy of prediction strongly depends on the number and variety of documents included in this training set.

### Classification methods based on citation

Far from controlled or uncontrolled descriptors, citation parameters have also been used for the translational classification task. Han *et al.* introduced a simple, comprehensive, and reproducible method to categorize whether a publication should be considered translational or non-translational research through the evaluation of bibliometric differences between both [25]. This study follows and expands the bibliometric approach of the previous lines of work, focusing on behavioral and social science research [26, 27]. Publication type is used as a proxy for study categorization and two types of translational research can be defined: primary or secondary translational research. On the one hand, a publication can be considered primary translational research if it is classified as an original clinical trial or clinical guidelines (as defined by the list of MeSH Pubtypes) [28]. On the other hand, articles that do not meet the previously described requirements, but are cited by primary translational research, are classified as secondary translational research. Although these studies are not translational themselves, they may still contribute to future

translational research. Using publication types and citations for categorization tasks may solve an overestimation that usually occurs when descriptor-based methods are applied [25].

Even novel bibliometric indicators were described as translational features of non-clinical research using citations and publication journals, such as the Translational Science score (TS score) [9]. Publication journals are categorized into four field categories: Clinical science (Clinical), Non-clinical science (Non-Clinical), Multidisciplinary (Multi), and Non-science and engineering (Non-S&E). The TS score is calculated using the share of clinical forward citations among all forward citations that a non-clinical publication has received. TS score is simply to replicate, relatively straightforward to calculate, offers good reliability and validity properties and can be implemented at scale. It is designed to capture the short-term use of basic research in subsequent clinical research, relative to its use in another basic research. This approach is consistent with Trochim *et al.*'s process marker model [29].

Even these translational predicting categorization methods were improved by supervised machine learning algorithms. The early reaction of the scientific community to an article, based on citation dynamics, provides enough information for a machine learning system to analyze it and predict its translation progress in biomedical research [12]. In this study, the

## Identification of biomedical translation

MeSH terms were used to classify articles as translational documents, like Weber's method [21]. However, the citation is used to track knowledge flow in translational research. The random forest algorithm, by means of machine learning, was used to calculate the probability that an article will end up being cited by a clinical publication [12]. Higher Approximate Potential to Translate (APT) scores represent documents that showed a higher probability of being cited by a clinical publication. In this milieu, documents with high APT scores could be considered as documents nearer to clinical application and, thus, more translational [12].

### Discussion

Translational research has become an important topic of interest for the biomedical community that is attracting the attention of the political sphere, responsible for deciding on research funding and grants [9, 25]. However, feasible methods to quantify or characterize the conditions for successful translation are still lacking. Therefore, to support the decision-making process, there is the need to develop a method to assess the position of research on the translational pathway. The identification of key points for the dissemination and application of biomedical research plays an important role in monitoring the progress of biomedical research and the flow of knowledge through the translational process, allowing us to assess the impact on public health and the effectiveness of the grant programs. Although it is possible to evaluate the translational progress of an individual work, there is no standard method to assess whether research areas, medical specialties, institutions or even nations are progressing on the bench-to-bed translational objectives [30]. For this reason, quantitative monitoring of this process is considered a troublesome task, and a special research effort should be focused on this purpose [25, 31].

Traditionally, an extended way to measure translational progress consists of analyzing how clinical articles cite basic science discoveries. It is noted that if the transformational discovery has a significant impact on human health, basic articles containing those discoveries tend to be cited by clinical articles. Nevertheless, there is a complex relationship between early scientific outcomes and the subsequent treatments to which they finally lead

[16, 32], and this way of measuring could undervalue the subset of basic biomedical research that lacks a direct link to clinical medicine [12].

Patents have also traditionally been used as a possible unit of analysis to measure translation. However, the use of patents presents some serious drawbacks, such as the time to know whether it will be used commercially and the subsequent uncertainty [9]. In contrast, articles need less time to publish the results obtained from a research study [33]. For this reason, although there have been papers that have linked research funding to their patent-related output [34], articles allow earlier analysis of the impact of translational research investment. In addition, bibliometric and citation analysis allows us to know how the knowledge generated is related and dispersed among authors, institutions, and even disciplines [35]. Several authors have developed methods to measure translation of publications using a bibliographic-based approach instead of patents.

The methods included in the present systematic review make up this set of methodologies attempting to identify and classify translation of the different levels of the biomedical literature. Unfortunately, there is no consensus on which one offers the best results [8, 23]. The choice of a method should be cautious and may depend on different factors and variables, such as the available data and information or the final purpose of the measurement. The combined use of several methods or indicators is recommended to provide more robust results [9, 22], as each method presents some strengths and limitations. In this context, a systematic review of the literature is necessary to bring together all the proposed methods and synthesize their particularities, advantages, and weaknesses. After critically reviewing the available literature, these methods can be divided into descriptor-based and citation-based classification methods.

One of the most relevant methodologies based on the use of controlled descriptors is the Triangle of Biomedicine [21], a method that has already been used by some authors to capture the degree of translation [36]. This graphical approach allows to visualize in a trilinear space the translational status of a set of documents



## Identification of biomedical translation

representing a research area, and even to track the evolution of a research area along the translational path over time. However, the use of PubMed Central (PMC) and Relative Citation Rate (RCR) as a measure of citation, although they can provide valuable additional information [37], may imply an underestimation compared to other databases, such as Web of Science or Scopus [21]. Other limitations of this method are those inherent to the use of the MeSH terms. These expert-assigned terms are not immediately available, as they take a while to be assigned to each article and may change in time as nodes and subtrees also change, splitting, merging, or introducing new descriptors. This may result in a percentage of documents that cannot be classified, as authors reported themselves [21]. In fact, several authors tried to solve some of the major drawbacks of this technique, assuming its relevance in the translational classification task [9, 12, 22].

As an alternative, uncontrolled descriptors can also be used. Although heterogeneous, these units of analysis may provide more complete and immediate information about the content of the research. In the past, experts in a specific field were required to manual curation of the later biomedical literature to track knowledge flow from bench to bedside. These kinds of analyses had a huge time and effort cost, especially, with large databases. With the recent development of algorithms and advances in the processing and availability of data, this challenge tends to disappear, giving the opportunity to measure and analyze the progress of biomedical research on a large scale [12, 22, 38]. However, this should be approached with caution, as experts might identify translational progress that is not reflected in the features that the machine learning systems work with [12]. Although metrics alone should not replace human supervision, this automated method of assessing translation can be an acceptable way to optimize the evaluation process [12]. Currently, several methods were proposed that belong to this group of machine learning-based methods [12, 23, 24]. Further work showed the usefulness of this type of approach using embedded words and how the use of custom embedded words increases the performance of this method [39].

Machine-learning algorithms tend to identify the translational status of articles with a high

level of discrimination and high accuracy [24]. Nevertheless, they could eventually become obsolete because of the change of classification over time, requiring a periodic update [12, 21, 23, 40]. Moreover, it is important to consider that the set of words featured in machine learning methods does not fully capture semantic or syntactic information. Also, supervised learning algorithms highly depend on a well-characterized training set. A lack of a well-documented corpus of each stage of the translational spectrum can cause difficulty in distinguishing between the different levels of translation [24]. In this sense, it can be assumed that these difficulties may be resolved in the near future.

On the other hand, other methodologies use citations as the unit of analysis of translational profile [9, 12, 25]. Some of these methods need a previous classification using publication types [25], journals [9] or descriptors [12], although the use of publication type may not be entirely accurate reflecting the true nature of the actual research conducted [25]. These methods seem to be more reliable and accurate, as they do not overestimate the number of translational research publications compared to MeSH terms-based methods [25].

Other advantages of citation-based methods, such as the TS Score, besides those mentioned above, are that they can apply to other research fields, can complement other metrics used for evaluating translational impact, and have a simple interpretation [9]. Even these methods present an acceptable sensitivity as they can obtain good reliability in documents with at least five citations. Nevertheless, the TS score depends on forwarding citations, a limitation shared by other forward citation-based measures, such as APT [12] or Han's secondary translational research measure [25], which could lead to a delay in the proper measure estimation [9].

Some of these citation-based methodologies use machine learning; this method allows one to perform a dynamic temporal evaluation of translational sciences, obtaining a graphical representation of these results. The APT score and other related indicators of all the publications available in PubMed can be calculated on the following website <https://icite.od.nih.gov/analysis> [12].

## Identification of biomedical translation

Overall, the methods discussed above try to classify publications into some predefined categories, performing qualitative analysis. However, quantitative approaches have also been described. In this sense, the Level Score was proposed to quantitatively evaluate articles and place them on the continuum spectrum of translational research, currently being the only method that provides this feature [22]. The limitations of this method are those of MeSH terms, previously mentioned. Also, when performing the citation analysis, there was a significant loss of articles, because of it uses a database for obtaining the publications (PubMed) and another for the citations (Web of Science) [22].

Although the present systematic review provides relevant insight into this emerging field, we need to name some limitations. The first shortcoming is related to the bibliographical databases. The results were conditioned by the search queries and the search engine that employed each database. Nevertheless, a detailed process is reported to make this study more reproducible. Another limitation is related to the information provided by the authors. Not all the documents reported similar information about the validity or reliability, making the comparison between methods difficult. Thus, it would be interesting to perform an experimental comparison of the different methods presented in this review applied to a standard set of documents. Furthermore, it is necessary to develop further translational studies that enable their adequate characterization, and consequently, support testing the performance of these methods more accurately.

In summary, there are few methods aimed at consistently measuring biomedical translation. Nonetheless, we found two groups clearly defined in our review, depending on whether they are based on descriptors or citations. All of them present limitations that need to be considered. Some of these limitations are related to the bibliographic databases or other time-dependent features (e.g., changes in MeSH terms or citation data). In this sense, there is still a lack of a gold-standard method for this purpose. The development of methods using information belonging to basic document metadata, instead of specific information provided by external databases, would be more consistent, as basic metadata are immediately avail-

able and remain constant. Undoubtedly, a method that reflects the accurate translational pathway of the different research developments will be crucial for policymakers and administrators to promote advances in science with application in daily clinical practice.

### Acknowledgements

This work has been supported by the Spanish State Research Agency through the project PID2019-105381GA-I00/AEI/10.13039/501100011033 (iScience), grant CTS-115 (Tissue Engineering Research Group, University of Granada) from Junta de Andalucía, Spain, a postdoctoral grant (RH-0145-2020) from the Andalusia Health System and with the EU FEDER ITI Grant for Cadiz Province PI-0032-2017. The present work is part of the Ph.D. thesis dissertation of Javier Padilla-Cabello.

### Disclosure of conflict of interest

None.

**Address correspondence to:** Dr. Antonio Santisteban-Espejo, Department of Pathology, Puerta del Mar University Hospital, Av. Ana de Viya, 21. 11009. Cadiz, Spain. Tel: +34-603470838; E-mail: antonio-santistebanespejo@gmail.com

### References

- [1] Wehling M. Translational medicine: can it really facilitate the transition of research “from bench to bedside”? *Eur J Clin Pharmacol* 2006; 62: 91-95.
- [2] National Center for Advancing Translational Science. Transforming Translational Science. 2019.
- [3] Austin CP. Opportunities and challenges in translational science. *Clin Transl Sci* 2021; 14: 1629-1647.
- [4] National Institutes of Health. National Center for Advancing Translational Sciences: about the CTSA program. 2021.
- [5] Llewellyn N, Carter DR, Rollins L and Nehl EJ. Charting the publication and citation impact of the NIH clinical and translational science awards (CTSA) program from 2006 through 2016. *Acad Med* 2018; 93: 1162-1170.
- [6] Medical Research Councils UKRI. Translational Research 2021.
- [7] EATRIS. About EATRIS European infrastructure for translational medicine.
- [8] Blümel C. Translational research in the science policy debate: a comparative analysis of documents. *Sci Public Policy* 2018; 45: 24-35.

## Identification of biomedical translation

- [9] Kim YH, Levine AD, Nehl EJ and Walsh JP. A bibliometric measure of translational science. *Scientometrics* 2020; 125: 2349-2382.
- [10] Feeney MK, Johnson T and Welch EW. Methods for identifying translational researchers. *Eval Heal Prof* 2014; 37: 3-18.
- [11] Dembe AE, Lynch MS, Gugiu PC and Jackson RD. The translational research impact scale: development, construct validity, and reliability testing. *Eval Heal Prof* 2014; 37: 50-70.
- [12] Hutchins BI, Davis MT, Meseroll RA and Santangelo M. Predicting translational progress in biomedical research. *PLoS Biol* 2019; 17: e3000416.
- [13] Kane C, Rubio D and Trochim W. Evaluating translational research. *Translational Research in Biomedicine* 2012; 3: 110-119.
- [14] Dodson SE, Kukic I, Scholl L, Pelfrey CM and Trochim WM. A protocol for retrospective translational science case studies of health interventions. *J Clin Transl Sci* 2021; 5: e22.
- [15] Grant J, Cottrell R, Fawcett G and Cluzeau F. Evaluating “payback” on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *Br Med J* 2000; 320: 1107-1111.
- [16] Williams RS, Lotia S, Holloway AK and Pico AR. From scientific discovery to cures: bright stars within a galaxy. *Cell* 2015; 163: 21-23.
- [17] Luke DA, Sarli CC, Suiter AM, Carothers BJ, Combs TB, Allen JL, Beers CE and Evanoff BA. The translational science benefits model: a new framework for assessing the health and societal benefits of clinical and translational sciences. *Clin Transl Sci* 2018; 11: 77-84.
- [18] Rollins L, Llewellyn N, Ngaiza M, Nehl E, Carter DR and Sands JM. Using the payback framework to evaluate the outcomes of pilot projects supported by the Georgia Clinical and Translational Science Alliance. *J Clin Transl Sci* 2020; 5: e48.
- [19] Moher D, Liberati A, Tetzlaff J and Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6: e1000097.
- [20] Ouzzani M, Hammady H, Fedorowicz Z and Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016; 5: 210.
- [21] Weber GM. Identifying translational science within the triangle of biomedicine. *J Transl Med* 2013; 11: 126.
- [22] Ke Q. Identifying translational science through embeddings of controlled vocabularies. *J Am Med Informatics Assoc* 2019; 26: 516-523.
- [23] Surkis A, Hogle JA, DiazGranados D, Hunt JD, Mazmanian PE, Connors E, Westaby K, Whipple EC, Adamus T, Mueller M and Aphinyanaphongs Y. Classifying publications from the clinical and translational science award program along the translational research spectrum: a machine learning approach. *J Transl Med* 2016; 14: 235.
- [24] Hsu YY, Clyne M, Wei CH, Khoury MJ and Lu Z. Using deep learning to identify translational research in genomic medicine beyond bench to bedside. *Database* 2019; 2019: baz010.
- [25] Han X, Williams SR and Zuckerman BL. A snapshot of translational research funded by the National Institutes of Health (NIH): a case study using behavioral and social science research awards and clinical and translational science awards funded publications. *PLoS One* 2018; 13: e0196545.
- [26] Rosas SR, Schouten JT, Cope MT and Kagan JM. Modeling the dissemination and uptake of clinical trials results. *Res Eval* 2013; 22: 179-186.
- [27] Schneider M, Kane CM, Rainwater J, Guerrero L, Tong G, Desai SR and Trochim W. Feasibility of common bibliometrics in evaluating translational science. *J Clin Transl Sci* 2017; 1: 45-52.
- [28] U.S. National Library of Medicine. Publication characteristics (publication types) with scope notes: 2020 MeSH PubTypes.
- [29] Trochim W, Kane C, Graham MJ and Pincus HA. Evaluating translational research: a process marker model. *Clin Transl Sci* 2011; 4: 153-162.
- [30] Smith C, Baveja R, Grieb T and Mashour GA. Toward a science of translational science. *J Clin Transl Sci* 2017; 1: 253-255.
- [31] Norris AE, Matsuda Y, Altares Sarik D and Pettigrew J. Implementation quality: implications for intervention and translational science. *J Nurs Scholarsh* 2019; 51: 205-213.
- [32] Urlings MJE, Duyx B, Swaen GMH, Bouter LM and Zeegers MP. Citation bias and other determinants of citation in biomedical research: findings from six citation networks. *J Clin Epidemiol* 2021; 132: 71-78.
- [33] Ihli MI. Acknowledgement lag and impact: domain differences in published research supported by the National Science Foundation. Master's Thesis, University of Tennessee; 2016.
- [34] Li D, Azoulay P and Sampat BN. The applied value of public investments in biomedical research. *Science* 2017; 356: 78-81.
- [35] Llewellyn N, Carter DR, DiazGranados D, Pelfrey C, Rollins L and Nehl EJ. Scope, influence, and interdisciplinary collaboration: the publication portfolio of the NIH clinical and translational science awards (CTSA) program from 2006 through 2017. *Eval Heal Prof* 2020; 43: 169-179.
- [36] Santangelo GM. Article-level assessment of influence and translation in biomedical research. *Mol Biol Cell* 2017; 28: 1401-1408.

## Identification of biomedical translation

- [37] Hutchins BI, Yuan X, Anderson JM and Santangelo GM. Relative citation ratio (RCR): a new metric that uses citation rates to measure influence at the article level. *PLoS Biol* 2016; 14: e1002541.
- [38] Tatonetti NP. Translational medicine in the age of big data. *Brief Bioinform* 2019; 20: 457-462.
- [39] Major V, Surkis A and Aphinyanaphongs Y. Utility of general and specific word embeddings for classifying translational stages of research. *AMIA Annu Symp Proc* 2018; 2018: 1405-1414.
- [40] Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow PM, Zietz M, Hoffman MM, Xie W, Rosen GL, Lengerich BJ, Israeli J, Lanchantin J, Woloszynek S, Carpenter AE, Shrikumar A, Xu J, Cofer EM, Lavender CA, Turaga SC, Alexandari AM, Lu Z, Harris DJ, Decaprio D, Qi Y, Kundaje A, Peng Y, Wiley LK, Segler MHS, Boca SM, Swamidass SJ, Huang A, Gitter A and Greene CS. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface* 2018; 15: 20170387.