Technological Forecasting & Social Change 163 (2021) 120490

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Multidimensional Scientometric indicators for the detection of emerging research topics

Haiyun Xu [a,b,*], Jos Winnink [c], Zenghui Yue [d], Huiling Zhang [e], Hongshen Pang [f]

[a] Business School, Shandong University of Technology, Zibo 255000, China
[b] Institute of Scientific and Technical Information of China (ISTIC), Beijing, 100038, China
[c] Centre for Science and Technology Studies (CWTS), Leiden University, Leiden 2300 AX, Netherlands
[d] School of Medical Information Engineering, Jining Medical University, Rizhao 276826, Shandong, China
[e] Taiyuan Library, Taiyuan 030024, Shanxi, China
[f] Library, Shenzhen University, Shenzhen 518060, Guangdong, China

ARTICLE INFO

Keywords:
Emerging topics
Frontier topics
Transformative topics
Scientometric indicators

ABSTRACT

The detection of emerging research topics (ERTs) is an important step in the promotion of potentially promising research. In this paper, we first distinguish the concept of ERT from that of common-related topics. We discuss multi-dimensional and practical bibliometric indicators to discover high-impact ERTs at a fine level of granularity. To this end, we focus on three objectives. First, we determine a method for identifying research topics from scholarly publications. Second, we developed a method for uncovering ERTs. Third, we conduct an experimental analysis to demonstrate the operationalization of our method. High-impact ERTs are selected based on the advice of field experts. Then, we divide the identified ERTs into seven patterns and propose different research and development (R&D) strategies for each topic. In general, previous studies have mainly focused on the attributes of "novelty" and "growth." Compared with extant ERT identification analysis, this study pays more attention to the future economic and social impact of ERTs and the reduction of uncertainty. This article is the first attempt to devise an ERT detection framework by comprehensively measuring its five characteristics (radical novelty, relatively fast growth, persistence and coherence, potential high impact, and uncertainty and ambiguity reduction) and proposing feasible operational indicators and processes. Thus, different R&D layout strategies for different economic and social-impact topics are outlined herein, which contribute significantly to the development of science and technology policies. The evidence presented suggests that our methodology can be used by decision makers to evaluate emerging scientific research.

1. Introduction

For research foundations and policymakers, the timely identification of emerging research topics (ERTs) is very important, as they help in stimulating the development of promising research topics. Emerging technologies have been the subject of much debate in academic research and are a central topic in policy discussions and initiatives to stimulate innovation. The growing number of science - and technology-related publications and the number of news items mentioning emerging technologies and research projects indicate the increasing attention received by emerging technologies and related topics. Various projects funded by the European Research Council (ERC) to identify emerging areas by topic such as the "Emerging Research Areas" and their Coverage by ERC-supported Projects (ERACEP) are examples of the increasing interest on this subject. The US Intelligence Advanced Research Projects Activity (IARPA) gives more importance to high-risk/high-payoff research, which is concerned with taking risks rather than going for quick wins or sure bets. With the mandate to generate revolutionary capabilities. In 2011, IARPA established its Foresight and Understanding from Scientific Exposition (FUSE) program, and one of the main purposes of this program is to nominate emerging technologies (Small et al., 2014).

A profound understanding and accurate recognition of the development and trends of ERTs play a crucial role in the deployment of successful innovation strategies and the optimal allocation of innovative resources. Therefore, policymakers and academic researchers are paying increasing attention to the identification of potential ERTs. ERTs are

* Corresponding author at: Business School, Shandong University of Technology, No. 88, Gong Qing Tuan West Road, Zibo 255000, China.
E-mail addresses: xuhaiyunnemo@gmail.com (H. Xu), winninkjj@cwts.leidenuniv.nl (J. Winnink), yzh66123@126.com (Z. Yue), lingkuailewuxian@sina.com (H. Zhang), phs@szu.edu.cn (H. Pang).

https://doi.org/10.1016/j.techfore.2020.120490
Received 12 March 2020; Received in revised form 22 November 2020; Accepted 24 November 2020
Available online 7 December 2020
0040-1625/© 2020 Elsevier Inc. All rights reserved.

considered to become increasingly crucial for the development of potentially promising research and R&D management policies. Moreover, it is important to be able to accurately recognize ERTs and their developing trend to support strategic R&D planning and to ensure the optimal allocation of innovation resources.

Although there has been an increase in the number of ERTs (Luo et al., 2018), little is known about the connotations and extensions of the concept of ERT and other related concepts such as research hotspots, frontier topics (FTs), and transformative topics (TTs). Good disambiguation will help us better understand the nature of an ERT. In prior studies, no clear boundaries were defined with regard to frontier or disruptive technologies. Additionally, although numerous studies have been conducted on ERTs, and various approaches have been proposed to detect them, there are major differences that remain among their connotations. It is possible that the challenge of identifying ERTs is partially caused by the lack of a clear definition for ERTs. Hence, the basic concept of ERTs is first defined to obtain a clear understanding of the connotation of ERT. Then, a more advanced and targeted identification method can be developed on the basis of this concept.

Upon closer examination of the literature, the detection of ERTs has shown some major shifts over the years: (1) from citation-based approaches to machine learning-based ones (e.g., Lee et al., 2018), (2) from measurement to identification (e.g., Poter et al., 2019; Wang, 2018), and (3) from papers to patents (e.g., Lee et al., 2018; Poter et al., 2019). Currently, judgment-based detection of ERTs has been primarily based on expert approaches (e.g., the Delphi method (Linstone and Turoff, 1975) and large-scale surveys (Xu et al., 2019)). There are two representative studies on ERTs: Rotolo et al. (2015) and Wang (2018). The former defines an ERT as follows:

…a radically novel and relatively fast-growing technology characterized by a certain degree of coherence persisting over time and with the potential to exert a considerable impact on the socioeconomic domain(s), which is observed in terms of the composition of actors, institutions, and patterns of interactions among those, along with the associated knowledge production processes. Its most prominent impact, however, lies in the future… so the emergence phase is still somewhat uncertain and ambiguous (P.13).

Rotolo et al. (2015) focused on societal and economic influences in addition to academic impact, whereas Wang (2018) focused only on ERTs in the academic field, providing the following description:

A radically novel and relatively fast-growing research topic is characterized by a certain degree of coherence and considerable scientific impact (P.293).

We argue that ERTs should not be confined to research topics. Their economic and societal effects should also be considered. We begin our study by considering the ERTs presented by Wang (2018). In addition to the academic impact of ERTs, we focus on their societal and economic impacts, similar to the approach followed by Rotolo et al. (2015). In detail, a set of multi-dimensional and practical bibliometric indicators is proposed to support the discovery potential of high-impact or transformative ERTs at a fine level of granularity and focus on ways to operationalize ERT criteria. To verify the feasibility and effectiveness of the method, the stem cell field was considered as a case study. Our method is expected to aid in making decisions about the prioritization of future scientific and technological research.

The remainder of this paper is organized as follows. Section 2 reviews the literature related to ERTs, identifies their relevant concepts, and analyzes the extant methods of identification. Section 3 presents a framework to identify ERTs, defines their criteria, and proposes a set of multi-dimensional and practical detection indicators. Section 4 empirically analyzes the ERTs of stem-cell research from 2004 to 2018. Finally, Section 5 provides the main conclusions of the study and directions for future research.

2. Theoretical background

Here, we first describe the ERT concept in SubSection 2.1, and discuss the scientific and technological innovations in SubSection 2.2. Then, various concepts of scientometrics are discussed in SubSection 2.3.

2.1. Defining ERT

Although there is some overlap between the concepts of emerging "technologies" and emerging "research topics," the concepts are distinct. It is important to note that technological advances are not necessarily based on scientific advances (Rotolo et al., 2015). Cozzens et al. (2010) described ERTs using terms such as rapid growth, newness, untapped market potential, and high-technology base. Rotolo et al. (2015) defined ERT using five key attributes: radical novelty, relatively fast growth, coherence, high impact, and uncertainty with ambiguity reduction. That is, they regarded an ERT as a radically new and relatively fast-growing technology that is characterized by a certain degree of persistent cohesion. An ERT has the potential to significantly impact socioeconomic domain(s). Such an impact can be described using a mixture of elements from the set of relevant actors, institutions, interaction patterns, and the process of associated knowledge production. We define emerging technology similar to the way Rotolo et al. (2015) defined it. First, the attribute of prominent impact must be reconsidered so that it remains relevant to the identification of ERTs. Rotolo et al. (2015) used this prominent impact to specifically refer to the societal and economic impacts of technologies. Meanwhile, Wang (2018) replaced the societal and economic impacts with scientific ones. The attributes of uncertainty and ambiguity are considered irrelevant. A rapidly developing research theme of radical novelty is characterized by its degree of coherence and scientific impact. Therefore, ERTs are assumed to exert prominent scientific impact within a short time period after their emergence. Burmaoglu et al. (2019) examined emergence discussions to understand the evolution of related concepts. They qualitatively reviewed the philosophy and complexity of science and several economic studies based on the emergence concept and found that it has mostly been used to describe unexpected and surprising events.

To enable a granular description of ERTs, we must first understand their evolution, laws, and trends as they pertain to specific fields.

2.2. Scientific innovation and innovation topics

In this study, innovation refers to both scientific and technological advances. Scientific innovation relies on the exploration of the laws of nature and the discovery of knowledge. Technological innovation focuses on the methods, means, and processes of transforming technology. Currently, science and technology are two driving forces behind an integrated innovation body in which they mutually promote or restrict each other instead of taking completely separate trajectories. (Chen, 2001; Xu et al., 2020).

Scientific competitive intelligence deals primarily with the content of scientific and technological innovation, mostly at the micro level or with fine granularity in the form of topics dominated by ERTs and frontiers of scientific innovation. These topics provide insights into specific fields and help identify hotspots. Therefore, the identification and prediction of innovation topics is crucial for the analysis of scientific and technological competitive intelligence.

In the literature, innovation topics are usually expressed using terms such as "topic identification," "topic finding," "label identification," "cluster labeling," and "category labeling" (Stein and Eissen, 2004). Although the term lacks a clear definition, "innovation topics" have been further described by Zhang et al. (2019). Their first definition states that "a topic is a collection of articles that are mathematically represented by their centroid, which is identified as the article sharing the highest similarity with all other articles on the topic." Their second definition

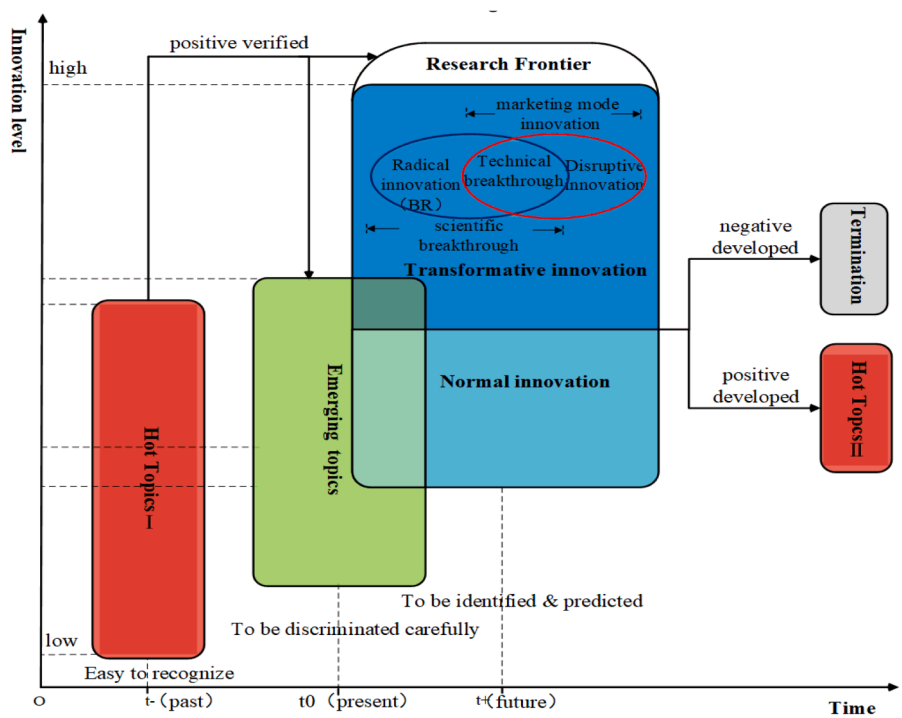

Fig. 1. Research frontier and transformative innovation-TI (Luo et al., 2018).

states that "a topic is geometrically represented as a circle, and its boundary is the largest Euclidean distance between its centroid and all articles." Historically, the identification of innovation topics is usually done using expert opinions. However, given the large amount of data that must be processed, manual identification becomes nearly impossible, and computerized algorithmic methods become more essential.

2.3. Similar concepts related to ERTs in scientometrics

Studies related to the concept of emergence can be traced back to the 1900s (Burmaoglu et al., 2019). Even though there are a growing number of publications regarding ERTs, the resulting concept remains ambiguous. There is still a lack of disambiguation between the concept

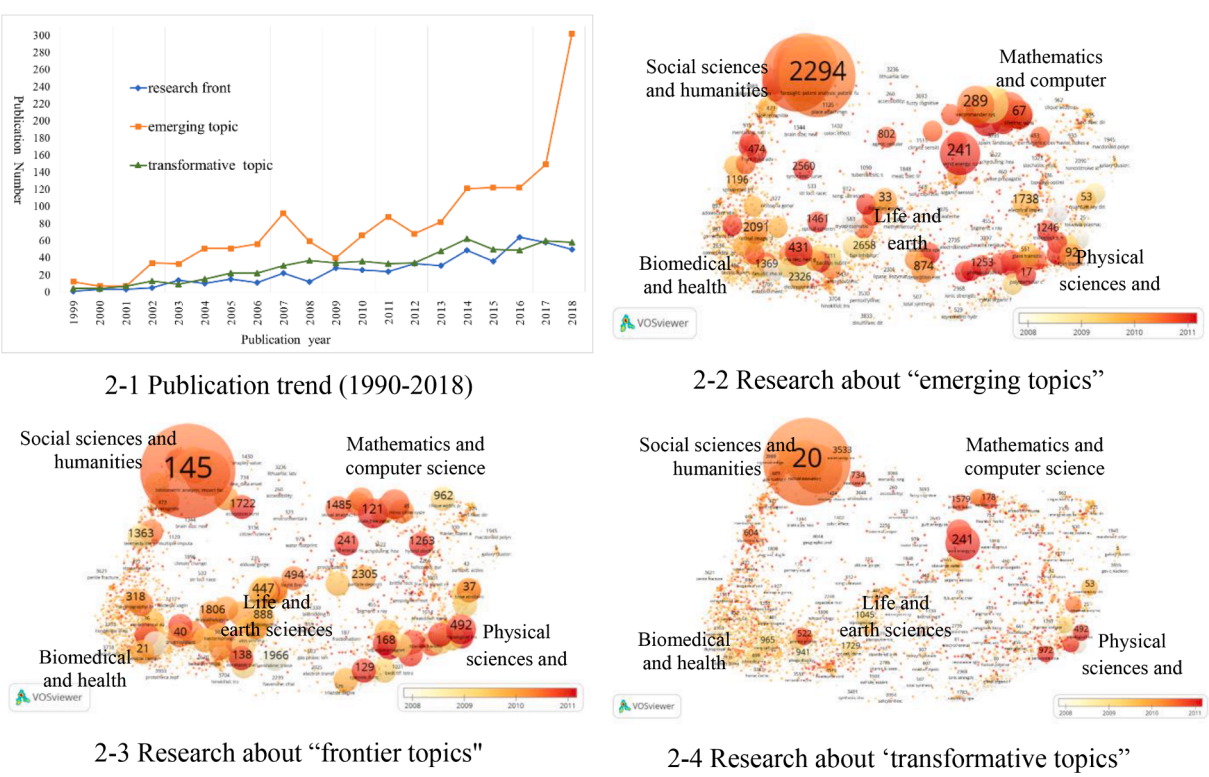

Fig. 2. Publication and overlay maps of "emerging topics", "frontier topics", and "transformative topics" (1990–2018) (each node is labelled with a topic identification number).

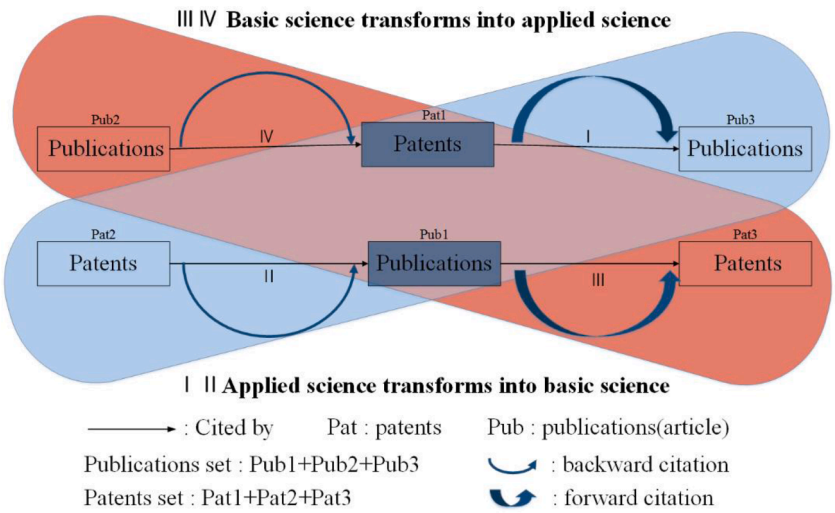

Fig. 3. Schematic of science–technology interaction.

of an ERT and related concepts such as FT or TT. The following discussion aims to provide clarity about these concepts. Xu et al. (2019) presented a number of phrases related to ERT such as "emerging topics," "research fronts," "emerging trends," "emerging technologies," and "emerging research fields." Although the meanings of these terms are not completely identical, they are used interchangeably in their study. Villaseñor Terán (2017) explained the differences and similarities of "emerging technologies" from "invention," "innovation," and "new and disruptive technologies" from a philosophical perspective. Based on his findings, he assessed "emerging technologies" and emphasized the dynamic behavior of technological development, while pointing toward concrete artifacts and procedures. Li et al. (2018) analyzed the interplay of technological emergence, disruption, and innovation using bibliometric methods and considered multiple theoretical definitions to model research embracing technological change processes, disruption, and emergence. Burmaoglu et al. (2019) performed a bibliometric study to strengthen the qualitative argument and found evidence of emergence in technological studies for comparison.

The differences among the aforementioned concepts related to ERTs lead to conceptual crossover and consideration of interdisciplinary approaches for technological (and commercial) emergence. To effectively identify ERTs, the concept developed by Luo et al. (2018), which focuses on frontier research ERT and its related concepts, is extended here after comprehensively reviewing the related work on technology prediction. Such disambiguation could help us better comprehend the nature of an ERT. Fig. 1 presents a schematic of the connotation of the relevant concepts considered in this paper. The two-dimensional coordinates are shown with O as the coordinate origin, the horizontal axis as the time dimension, and the vertical axis as the innovative dimension. The innovative axis, which ranges from incremental to radical, covers a wide spectrum. On the time axis, $t_0$denotes the observation time point, and t_0 divides the entire time axis into three periods: past, present, and future. From the perspective of time, "research frontier"' refers to future innovations. Compared with the "research hotspots," "research frontiers" and "emerging research" are more innovative. In terms of the degree of innovation, scientific and technological innovation can be further divided into "gradual innovation" with a low rate of innovation and "transformation innovation" with a high rate of innovation.

Although this study focuses on the operationalization of the discovery of ERTs, the discovery of FTs and TTs is also included in this work, as these concepts partially overlap. For example, TTs can be regarded as ERTs with a high level of innovation. Research on ERT, FT, and TT exhibited different trends over time (Fig 2-1). The concepts of TT and ERT existed before that of FT. Overall, the attention received by all three types of topics has increased every year. However, interest in ERTs has grown considerably in recent years, although it has always been substantially greater than the interest in TTs and FTs.

To identify the research activities, we refer to the overlay map

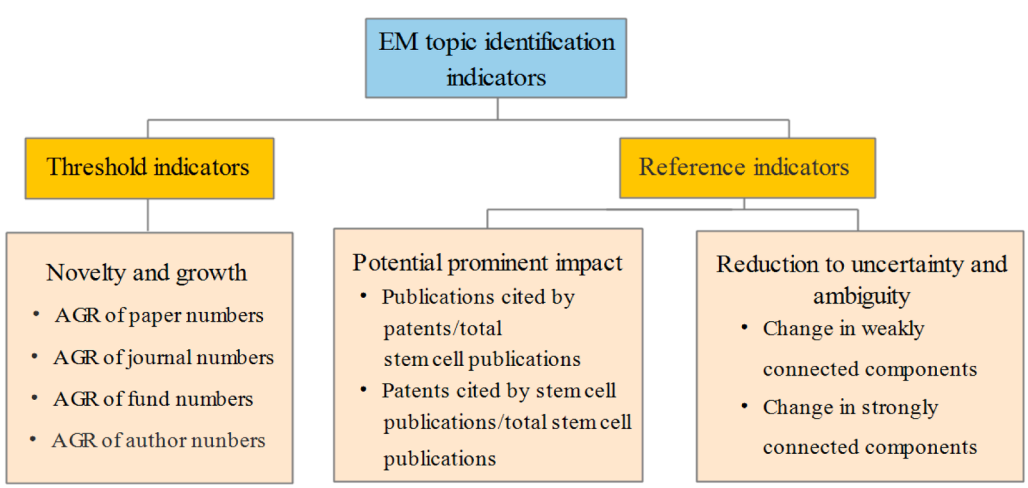

Fig. 4. ERT identification indicators.

Table 1
Concept disambiguation of "TTs," "ERTs," and "FTs".

Concepts	*Focus aspects*	*Keywords*
transformative topics (TT)	method	network, model, case study
	analysis object	literature, patent
	Focus	**innovation**, firm, radical innovation, new product, customer, company, challenge, change
emerging topics (ERT)	method	model, road mapping, bibliometric, case study
	analysis object	patent, publication
	Focus	**technology**, innovation, product
frontier topics (FT)	method	mapping, citation, bibliographic coupling
	analysis object	journal, article, paper, publication, patent
	Focus	**research**, research frontier, knowledge structure

provided by *VOSviewer* software (Eck and Waltman, 2019). The search strategy was extended to publication titles, abstracts, and keywords; however, they were sourced only from journals that highlight innovative concepts on the topics of scientific policy, management, and scientometrics.

In these maps, the items are colored according to the classification base map produced using the Leiden algorithm. There are five classifications on the base map: "social sciences and humanities," "mathematics and computer science," "physical sciences and engineering," "life and earth sciences," and "biomedical and health sciences." A shorter distance between topics indicates a stronger relation. In the overlay maps of ERT, FT, and TT (Figs 2-2, Figure 2, and Figure 2, respectively), each node represents a single topic (e.g., node 1 represents topic 1) and a larger node size means more publications belong to that subject (the numbers are identifiers; they have no relation to the number of publications). The different colors indicate different time slices: orange and yellow represent topics existing before 2010, and dark red represents topics created after 2010. Notably, there is no clear distinction between the ERT, FT, and TT maps, and they all converge at the "social sciences and humanities"' region.

To analyze the concepts of ERTs, FTs, and TTs, and to highlight the focus of each type, the text content that accounts for the largest proportion of each concept in the overlay maps was selected for detailed comparison. Here, topic 2294 from the ERTs, topic 145 from the FTs, and topic 20 from the TTs were selected. The titles and abstracts of the papers on these three topics were selected for further analysis by text processing, and the "Wordcloud" package of Python was used to extract meaningful phrases and their frequency. Finally, the 50 most frequent words and phrases that were used as the original keywords were selected and manually refined for further concept discrimination. Table 1 lists the keywords that best represent the respective concepts.

First, from the perspective of the methods employed for research on each topic, we find that FTs are researched mostly using traditional bibliometric methods, while TTs are not. Both traditional bibliometric methods, model, and road mapping methods are used in the research of ERTs. Second, the object of analysis is different. The research on TTs emphasizes case studies, while the research on FTs focuses more on scientific articles than patent documents; however, the research on ERTs comprises all types of publications. Finally, from the perspective of research focus, the studies on TTs focus more on topics that are directly related to the real-world economy, such as products, companies, economies, markets, users, and technological change. In contrast, studies on ERTs emphasize the prediction of future directions for development (e.g., the development of the technology industry) as well as the frameworks and strategies for identifying innovation. Studies on FTs mainly focus on the development of scientific research, such as the development of knowledge structures.

In this study, we compared TTs with ERTs because TTs are often closely correlated with ERTs. However, an ERT may or may not be a TT. In the following study, we will focus on multidimensional scientometric indicators for ERT detection, not TTs, without any further exploration as to whether some of the discovered ERTs are TTs. We will conduct such research in the future.

2.4. Scientometric indicators for ERT detection

To measure emergence, various approaches have been proposed in the literature (De Haan, 2006; Goldstein, 1999). Cozzens et al. (2010) noted that most prior studies concentrated on measuring ERTs rather than identifying them. Since then, many researchers have explored various approaches for ERT identification. The perspectives that were found range from the measurement of macro-level players to micro-level scientific topics (Small et al., 2014). The data and methods range from expert opinions (Porter et al., 2002) to the bibliometric analysis of R&D publications and patents (Glänzel and Thijs, 2012). In these methodologies, various scientometric variables such as novelty and growth serve as core modules (Small et al., 2014). Scientometric techniques are intrinsically more effective for retrospective analysis than contemporary examinations (Rotolo et al., 2015).

Small et al. (2014) combined direct citation and co-citation analyses to identify ERTs in terms of novelty and growth indicators. Small et al. (2014) noted that lexical-based methods could list emerging topics within a specific field, but that they were not suitable for identifying ERTs across a wider scope. Meanwhile, the results revealed by IARPA suggested that the clues in the wording and relationship among scientific papers and patents could predict the success of a given research (Reardon, 2014). Owing to the time lag associated with citation-based methods, there has been an increase in the use of lexical-based approaches (Xu et al., 2019). Rotolo et al. (2015) divided scientometric methods into five main categories: indicators and trend analysis, citation analysis (including direct citation, co-citation analysis, and bibliographic coupling), co-word analysis, overlay mapping, and hybrid approaches that combine two or more of these categories. Rotolo et al. (2015) discussed approaches developed in other fields, such as science and technology policy and management.

Referring to Rotolo's classification, this article divides ERT recognition methods into three categories based on the method used: citation-based, lexical-based, and hybrid approaches.

2.4.1. Citation-based methods

Citation-based methods can be further divided into three sub-groups: direct-citation network analysis (Waltman and van Eck, 2012), co-citation network analysis (Small, 1973; Small and Griffith, 1974), and bibliographic coupling network analysis (Glȧnzel & Thijs, 2012; Huang and Chang, 2014). Iwami et al. (2014) analyzed the time transitions of centrality measures to detect the leading paper among many papers. Jarić et al. (2014) observed how the emerging research subfields were characterized by the comparably more recent references (which had a higher proportion) published in the preceding two years. Santa Soriano et al. (2018) analyzed citation patterns and co-occurrence keywords while assessing their impact and maturity to propose future research ideas related to technological observatories. Wang (2018) proposed a series of indicators for identifying ERTs, and these indicators were constructed based on the direct citation relations of individual publications. This is considered to be a flexible approach because researchers interested in identifying ERTs in a particular field can reset the current parameter values based on the field under investigation.

2.4.2. Lexical-based methods

Porter and Detampel (1995) counted the keywords in publication abstracts and trend analyses based on Fisher–Pry curves to identify ERTs. Lee (2008) used co-word analysis to measure the latest research trends and aspects underlying the scientific and technological

documents used by researchers and policy planners. Takeda and Kajikawa (2009) developed a topological clustering method to detect ERTs. Roche et al. (2010) combined terminological evolution and diachronic clustering analysis to identify promising emerging technologies in the field of molecular biology. Tu and Seng (2012) proposed novelty and published volume indices for emerging topic detection. The concepts of terms, topics, and candidate research topics were used to investigate ERTs. Arora et al. (2013) conducted a study on website text mining to probe ERTs. Hu et al. (2013) analyzed a co-word network to reveal the status and trends of library information and systems. Hansmann and Niemeyer (2014), Yan (2014), and Liu et al. (2015) excavated and analyzed topics, keywords, and controlled terms in scientific and patent literature using natural language processing to identify and explore ERTs. Weismayer and Pezenka (2017) performed an automated longitudinal latent semantic analysis of keywords to determine ERTs. Joung and Kim (2017) proposed a keyword-based patent-analysis model to monitor ERTs. Li (2018) proposed the use of keyword aggregation across topics to identify ERTs by analyzing dispersed risk signals, which proved to be a timely and efficient method. Carley et al. (2018) and Porter et al. (2019) proposed an indicator set that operationalized the four aspects denoted in the FUSE-based and Rotolo et al.'s (2015) descriptions of technical emergence.

2.4.3. Citation and lexical hybrid approaches

Chen (2006) developed CiteSpace II using co-citation analysis and burst detection to detect and visualize emerging trends and radical changes in scientific disciplines over time. Guo et al. (2011) presented a mixed model that combined the frequency of specific words and changes in the interdisciplinarity of cited references to predict ERTs. Zitt et al. (2011) examined the convergence of citation- and word-based thematic mapping approaches. Huang et al. (2017) introduced a systematic approach that applied co-classification analysis to reveal the technical evolution process of a certain technical field, co-word analysis to extract implicit or unknown patterns and topics, and main path analysis to discover significant clues about technology hotspots and development prospects. Xu et al. (2019) used several machine learning models to detect and predict ERTs. Specifically, thematic structures were first extracted from scientific publications using the dynamic influence model (DIM) (Gerrish and Blei, 2010). Then, the growth, coherence, influence, and novelty indicators were calculated. The first three indicators were based on the DIM model, and the last one was based on the citation.

2.5. Framework for ERT detection

Our framework for discovering ERTs consists of the following two steps: finding meaningful topics and identifying which of these topics are ERTs in terms of scientometric indicators. We first provide a brief review of meaningful topic discovery and scientometric indicators pertaining to ERT detection.

2.5.1. Discovering meaningful topics

Meaningful topics are usually discovered from the citation network of a target field (i.e., knowledge network). Community detection algorithms are often used to understand the structure of large and complex networks (Xu et al., 2018). It has been well documented that different community detection algorithms generate different thematic structures, and no universal testing standards are available for evaluating their results.

The taxonomy created by any particular process can contain artifacts derived from the process (Leydesdorff, 1987). Therefore, the justification for the effectiveness of the proposed methodology requires considerable discussion. Klavans and Boyack (2017) stated that in most fields of science, accuracy is of paramount concern. Admittedly, several fields lend themselves more accurately than other fields. This is particularly true for fields wherein physical properties can be measured, gold standards exist, and a great deal of research is replicated. However, none of these conditions are met when it comes to the delineation of topics or the creation of taxonomies in the scientific literature.

No single or best solution for topic extraction can be obtained from the same dataset. There are two main reasons for this. First, the structure of a body of knowledge varies from one individual to another with regard to perception. That is, more than one valid theme can be constructed depending on the perspective that is applied. Second, topical structures are reconstructed for specific purposes, notwithstanding the existence of a "best" solution for a given purpose (Boyack et al., 2017).

2.5.2. Multidimensional indicators for ERT detection

Next, a research topic can be considered to be an ERT if it displays the characteristics of five attributes: novelty, growth, persistence and coherence, potential high impact (i.e., societal and economic influence), and uncertainty and ambiguity reduction.

2.5.2.1. Novelty, growth, persistence, and coherence. In innovation studies, novel ideas are a result of combining existing pieces of knowledge in an original and unprecedented way (Schumpeter, 1939).Hence, the novelty of an ERT may take the form of discontinuous innovations derived from radical ones (Small et al., 2014). Novelty is clearly central to ERTs (Rotolo et al., 2015). Bai et al. (2017) proposed that if a new topic has been continuously reproduced over more than two reference time windows in recent years, the knowledge topic might represent the germination of an ERT, and it can be considered a candidate. Mairesse and Pezzoni (2018) formulated novelty as a novel combination of referenced journals appearing for the first time in an article. Thus, a novel article reports a new combination of journals in its reference list. It is assumed that an ERT embeds a clear degree of novelty in the technology that sets it apart from the status quo (Li et al., 2018).

The most common method of evaluating novelty is to compare the similarity of a candidate topic with its prior counterparts. Liu (2007) vectorized patent texts and evaluated their novelty by calculating their vector distances. Hido et al. (2012) represented a patent document as a bag of words and evaluated novelty by calculating the sum of the "word age" of vital words. Qi (2015) proposed three indices: "absolute novelty, " "relative novelty," and "combined novelty," which can be used for patent-novelty analysis with automatic semantic text analysis.

The second attribute of an ERT is "fast growth" or, at least, "growth" (Small et al., 2014).Growth may be measured from multiple perspectives, such as the number of scientists involved, funding bodies, or knowledge outputs (e.g., academic publications and patents). Further, growth can be modeled by fitting logistic or exponential curves to a time series to project future trends (Roper et al., 2011). For the radical novelty attribute, the fast growth of a technology must be contextualized. Boyack et al. (2014) summarized two patterns of ERTs. In one type, it is not focused in the beginning, but later on, it grows explosively. The other type rapidly becomes a focus area and grows steadily. Therefore, one technology may grow more rapidly than other technologies in the same domain(s). In this case, the term "relatively fast growth" may be more suitable.

The third attribute of an ERT is "persistence and coherence." ERTs exhibiting persistent trends can ultimately gain economic and societal influence, and such topics tend to have higher merit. The coherence of an ERT is described as the "convergence of previously separated research streams" that persist over time (Day and Schoemaker, 2006), "convergence in technologies" (Srinivasan, 2008), and technologies that "have already moved beyond the purely conceptual stage" (Stahl, 2011). Coherence refers to the internal colloquial characteristics of a group (e.g., as "sticking together", "being united", "logical interconnection," and "congruity").

We argue that "persistence" is the premise of coherence, which is more important than persistence when identifying ERTs. If a topic continues to develop and, in the meantime, the coherence increases,

which indicates that the topic becomes more distinctive (i.e., identifiable) and that the core research becomes more coherent.

Although the persistence and coherence of an ERT are not directly related to its impact, the former refers only to its duration. There may be several reasons why a topic is persistent. For example, it may have a significant socioeconomic impact in the future. This is because a research topic needs to be continuously exploited to finally result in influential scientific discovery. In particular, it is a long process that starts with scientific research and extends to technological development. In the end, it exerts its economic and societal influence. During this long period, scientific problems are continuously conquered, technological innovation is promoted, and industrial development takes place.

2.5.2.2. High impact. ERTs potentially "yield benefits for a wide range of sectors" (Martin, 1995), "create a new industry or transform existing ones" (Day and Schoemaker, 2006), "exert much enhanced economic influence" (Porter et al., 2002), or change "the basis of competition" (Hung and Chu, 2006). ERTs may exert a widespread impact by spreading through multiple levels of the socioeconomic system, such as organizations and institutions as well as knowledge production processes and technological regimes. One important reason policymakers pay considerable attention to ERTs is their potential and immense capacity to affect the entire socioeconomic system. Accordingly, we identified "high impact" as another key attribute of ERTs.

Mutual citations among papers and patents are important for the spread of knowledge between science and technology. In accordance with the previous studies discussed below, we argue that the citation from patents to research papers and vice versa could represent the potential of a candidate ERT. Narin and Noma (1985) and Schmoch (1993) analyzed patent citations and found several types of interactions between science and technology. Additionally, the strength of the interaction varied depending on the field. Meyer et al. (2010) established a connection between subject areas and technical fields using patent citations and found that applied science with high technological relevance had considerably more influence (Meyer and Debackere, 2010). Other researchers have used two-way citations to investigate the relationship between science and technology, owing to the deficiency of the one-way analysis. Gao et al. (2012) analyzed the interaction between science and technology using a paper–patent hybrid co-citation analysis. They also used this method to study the mutual fusion of science and technology in knowledge flow. Huang et al. (2015) found that the cross-citation frequency of science and technology in the field of fuel cell research was low but gradually increasing, indicating an enhanced trend of correlation and integration. Kwon et al. (2017) proposed a method for identifying the unintended consequences of emerging technologies. They employed the latent semantic analysis technique to yield multiple groups of contextually similar terms from future-oriented data sources, comprising the concerns of both experts and the public regarding future technologies.

Scientific and scholarly research may result in novel discoveries. The nature and impact of such a discovery on the cognitive structure and evolution of science may vary considerably. The impact of discoveries may extend beyond the domain of science and may be critical for technological applications, and even to innovations and products. Scientific discoveries and their incorporation in technology are often interlinked in complex ways within research and development (R&D) systems. It is possible to detect upcoming emerging technologies at an early stage based on the evolution of science–technology linkages (Winnink, 2018).

2.5.2.3. Reduction to uncertainty and ambiguity. The evolution of ERT technologies is always accompanied by uncertainty. As these technologies are innately still under development, their final influence can only be known in the future. Although it may be unintended and undesirable, failures are inevitable, especially in high-risk research. The non-linear and multi-factor nature of ERTs comes with a certain degree of autonomy, which in turn makes it challenging to predict future outcomes (De Haan, 2006; Mitchell, 2007). Consequently, understanding the probabilities associated with each possible outcome may be particularly problematic (Stirling, 2007). The level of impact of an ERT needs to be measured, and it becomes more certain as time progresses. Along with uncertainty in its emergent and growth phase, the most accentuated characteristics of an ERT are its potential for strong economic influence and the uncertainty and ambiguity of this influence (Hung and Chu, 2006).

The long-term impact of serendipitous discoveries also makes the measurement of serendipity challenging. Hence, uncertainty is an important feature that cannot be neglected, especially in high-risk R&D, where failures are inevitable (otherwise, identifying ERTs would be straightforward). Yaqub (2018) discussed some conceptual difficulties related to the unit of analysis and potential sources of over- and under-estimation of serendipity and concluded that serendipity measurement calls for qualitative efforts. However, to our knowledge, the evaluation of uncertainty and ambiguity remains largely unexplored in scientometric studies. It has been shown that numerous FT topics initially occur in the form of weak ties among disciplines (Wei et al., 2017), which is an initial signal of interdisciplinary and technological integration. Therefore, the variation of strong and weak ties could be used to capture the growth of ERTs (Xu, Hao, An, Yang& Wang, 2019). Min et al. (2019) proposed a citing-structure perspective for observing the unfolding of breakthroughs based on the variations of a knowledge structure. The results showed that compared with mediocre papers, breakthrough papers show outstanding performance in terms of the number of nodes, number of edges, average clustering coefficient, number of components, and average degree. However, indicators such as network density, maximum betweenness centrality, and maximum closeness centrality are not significantly discriminative.

We suppose that although there is a significant amount of uncertainty in candidate ERTs, which is difficult to measure, this uncertainty is not random. The uncertainty of the development of ERTs is due to the rare amount of data that can be leveraged for the predictive analysis of weak signals. The existing literature does not clearly define the concept and characteristics of weak signals of ERTs, nor does it clearly show which information belongs to the category of weak signals, leading to a vague understanding of weak signals of ERTs. Therefore, the analysis of weak signals is often used as an additional subject of interest in the analysis of current technology foresight and serves as a simple signal-assisted monitoring method that has not yet been fully explored. We argue that weak ties in the evolution network of scientific innovation are a type of weak signal. During the alternation between new and old innovations, there are many weak linkages between old and new knowledge systems. Therefore, it is difficult to identify the development trend, and the development of multiple ERTs seems to be full of randomness; therefore, it is difficult to effectively understand the changing trends of weak ties.

Therefore, we suppose that the development trend of ERTs can be determined based on its position and changing trend in the entire innovative network. To this end, an ERT is assumed to cause significant changes in the structure of a knowledge network. At the emergence stage of an ERT, owing to its novelty and other factors, a strong tie with other topics does not emerge. As the ERT grows and changes, a strongly connected community is gradually formed around the ERT.

2.5.3. Research objective

This study aims to provide and test a method to identify ERTs within a specific domain at a fine topic granularity. The method focuses on the operationalization of multidimensional scientometric criteria, including the uncertainty and ambiguity of an ERT. By focusing on a specific domain, we can obtain more specific and comprehensive information about the targeted area of study. Porter et al. (2019) conducted a similar study in which a sufficient number of discriminatory terms and their

relations were recognized for each ERT using topic clustering and other complicated methods. However, these black-box methods implicitly introduce unexpected errors and uncertainty (Velden et al., 2017). Hence, unlike that of Porter et al. (2019), our study focuses on directly identifying potential ERTs. This work also differs from that of Wang (2018), who considered all science domains simultaneously. In our opinion, only one domain should be taken into consideration at a time. Thus, more specific and comprehensive information can be obtained about the targeted domain. Another significant distinction is the attribute of uncertainty and ambiguousness, which is of little relevance in the study of Wang (2018).

3. Methodology

3.1. Topic discovery

In this study, we aim to identify micro-level ERTs and employ citation and lexical hybrid approaches. The citation method was used for topic identification, while the lexical method was used for topic labeling. The Leiden algorithm (Waltman and Van Eck, 2012) is one of the most cutting-edge algorithms used to discover community structure in (very) large datasets. The current updated version of the Leiden algorithm uses a smart local-moving (SLM) algorithm to optimize speed (Traag et al., 2018; Waltman and Van Eck, 2013). In this study, we used the Leiden algorithm to create a set of meaningful topics leveraging the entire Web of Science (WoS) database. The Leiden algorithm relies on direct citation relations and works at the level of individual publications, facilitating the accurate discovery of a topic. Hence, a three-layer hierarchical classification of all publications in the database was generated. These system levels are as follows: "macro", which consists of 20 research areas; "*meso*", which consists of672 research areas; and "micro", which consists of 4047 research areas. Using the micro-level of classification, we identified relevant publications in the stem cell field.

3.2. Multidimensional indicators for ERT detection

In the next three subsections, we define and operationalize eight scientometric indicators to instantiate the multidimensional indicators for ERT detection.

3.2.1. Novelty, growth, persistence, and coherence as attributes and evaluation indicators of ERTs

The novelty and growth indicator measures the novelty of a candidate topic from different aspects, including four secondary indicators: the average growth rates (AGRs) of the number of papers, journals, funds, and authors. Both the number of papers and the number of journals represent the research attention of the topic; the number of funds represents the funding status of the topic, which emphasizes the degree of strength of R&D input, and the number of authors represents the degree of attention paid to a topic by scholars.

Step 1(Novelty and growth): The AGR of the number of papers for each topic over the entire time interval i is obtained by calculating the growth rate of the number of papers for each topic in each interval. This allows for the identification of the time at which each topic appeared. Subsequently, more weight is given to recently appearing topics. Topics that disappeared after appearing previously are treated as less novel and excluded from further analysis. The AGR can be defined as follows:

$$\overline{V} = \frac{1}{N_T}\sum_{t-1}^{t+i}\frac{N_{t+i}-N_t}{N_t} \tag{1}$$

Here, $\overline{V}$represents the AGR of a candidate ERT, t represents the interval sequence, i represents the sequence number of intervals, and the adjacent interval sequence is represented by T ($T = i$-1). N_T denotes the number of adjacent time intervals involved in the AGRs during calculation. Each growth rate calculation involves two adjacent time intervals represented by $V_T.N_t$ represents the number of topic entities in the ith interval sequence, including publications, journals, authors, and funding bodies related to the topics considered in this study. The AGRs are calculated for all the intervals with respect to the number of journals, funds, and authors for each topic.

Step 2 (Persistence and coherence): Here, we propose the use of the visualization method, a topic evolution map that evaluates the persistence and coherence of candidate ERTs. The evolution of a candidate ERT refers to the development of an innovation topic and reflects the emergence, diffusion, and advancements of a topic. The evolution path of an ERT can help determine its persistence and coherence. In this study, we use the topic relevance between two adjacent time slices as the coherence degree value. Jaccard similarity is used to calculate the topic relevance between different time intervals. The numerator of the Jaccard similarity is equal to the number of publications shared by adjacent time intervals, while the denominator is the sum of the number of publications from both time intervals. The relationship between two different topics is proportional to Jaccard similarity. If the Jaccard similarity continues to increase, the coherence degree of the candidate ERTs is assumed to be gradually increasing.

3.2.2. Evaluation indicators of potentially high impact ERTs

Scientific and technological innovation should lead to the promotion of economic development through the application of various processes. The publications cited by patents represent the ability of basic science to convert to applied science. The patents cited by publications represent the ways in which applied science has the ability to support basic science. A higher value of these indicators represents a greater potential impact.

3.2.2.4. Evaluation theory of linkage between science and technology. Theoretically, exploring the scientific and technological relations in a specific field through the citation analysis of papers and patents requires a comprehensive analysis of the mutual citation of papers and patents. Specifically, the first step is to obtain the basic data: the target papers and their citations in a specific field, the target patents, and their references in a specific field. The second step is to perform a citation analysis of the target literature and its citation data.

Fig. 3illustrates the type of analysis that needs to be performed and shows a comprehensive study of the interrelated mechanism of scientific and technological innovation through the mutual citation of papers and patents. In Fig. 3 targeted patents (Pat1) and targeted papers (Pub1) are represented as the centered objects of analysis. Subsequently, through the citation linkage from or to the targeted analysis objects, four citation paths (I, II, III, and IV) are formed. Along the timeline, four paths evolve from left to right. Paths II and IV indicate backward citation, and paths I and III indicate forward citation.

When taking the patent (Pat1) as the starting point, two horizontal paths (I and IV in Fig. 3) are formed. Path I represents the citation to the target patent (Pat1) from the post-scientific paper (pub3), and path IV indicates the citation from the target patent (Pat1) to the prior scientific papers (Pub2). Taking the scholarly publication (Pub1) as the starting point, two horizontal paths (II and III) are formed. Path II indicates the citation to the prior patent (Pat2) from the target paper (Pub1), and path III indicates the later citation from the patent (Pat3) to the target scientific paper (Pub1). The red shaded region in Fig. 3 (top left to bottom right), including path III and path IV, indicate the citations to the papers from the patents. The blue shaded region (bottom left to top right), including path I and path II, indicates the citations to patents from the papers.

Among the four citation paths, path I and path II are the advancement of basic science *paths* representing the advancement of basic science by applied science. Path III and Path IV are *technology innovation paths* representing the transformation of basic science into applied science. To fully exploit the topic linkages between science and technology,

it is necessary to analyze the path of transmission between science and technology.

3.2.2.5. Evaluation practice of linkage between science and technology. In practice, when analyzing the linkages between science and technology, the existing approaches mainly focus on patents cited by papers, that is, paths II and IV of the evolution path (backward citation), and in most cases, path IV is the only choice. Our method differs from the existing ones in that we consider paths I and III, which means that the papers that cite patents and the patents that cite papers are also considered. Therefore, in our method, the entire scientific discovery path is used, which includes the exploration of paths I and II for a more accurate analysis of the scientific and technological associations of ERT topics. An analysis of all four paths of the citation relationships among papers and patents makes the analysis of the linkage in science and technology more comprehensive and accurate.

The potential for prominent impact (societal and economic influence) indicator includes two secondary indicators: publications cited by patents/total number of publications on a topic (U_{BA}) and patents cited by publications/total number of publications on a topic (U_{AB}). The connection between basic science and applied science highlights the potential influence of the topic on society or the economy.

In this study, we calculate two main indicators, U_{BA} and U_{AB}, to evaluate the potential for high impact. The indicator U_{BA} represents the ability of basic science to transform into applied science, while the indicator U_{AB} represents the ability of applied science to transform into basic science. These indicators are calculated as follows:

U_{BA} = publications cited by patents/total number of publications on a topic, where publications cited by patents = path III + path IV;

U_{AB} = patents cited by publications/total number of publications on a topic, where patents cited by publications = path I + path II.

3.2.3. Evaluating the evolution of uncertainty and ambiguity in ERTs

This study intends to evaluate or predict the possibility of the future development or decline of ERTs based on the evolution of weak ties in ERTs. Based on the changes in weakly and strongly connected communities, the position and structure of ERTs in the entire knowledge network are analyzed to predict the developmental trends of ERTs in the future. The rationality is described as follows:

3.2.3.6. Breakthrough innovation will affect knowledge structure. From the perspective of network dynamics, the inclusion of new information will affect the stability of the knowledge network structure (Chen, 2012). Therefore, the impact of knowledge innovation on original knowledge can be identified by measuring the degree of network structure change. Dahlin and Behrens (2005) used the structural differences in patent citation networks to identify disruptive innovations and verified them with patents related to *tennis rackets*. The results showed that the greater the difference, the higher the probability of breakthrough innovation. Although the "structure identification" of the network is not equivalent to "topic identification" and "structure change" does not represent "topic change", this change can, to some extent, characterize the change in the research topic.

3.2.3.7. Change of strong–weak tie in the network. Weak ties are an early feature of interdisciplinary and technological integration. In the network structure, a weak relationship refers to a type of relationship where the strength of the node relationship is lower than a threshold, and it is opposite to "strong tie." According to Granovetter (1983), strong ties maintain relations within organizations, while weak ties are the links between groups and organizations for information transfer, and weak relationships can enable a larger network to achieve structural cohesion. Rhee (2004) has empirically demonstrated that the weak tie of a network is conducive to breaking the barriers of self-roles and ensuring the diversity of network members, and is more likely to produce brainstorming effects in communication. Zhang (2011) detected the evolution of weak-tie data in the network from the perspective of the overall network and individual networks and identified possible development frontiers. Wei et al. (2017) identified the frontier interdisciplinary topics of information science through weak linkage analysis in co-word networks.

We identify and detect ERTs by observing the changes in the knowledge structure. Herein, the reduction of uncertainty and ambiguity indicators involves three secondary indicators: the amount of change in three types of knowledge networks, weakly connected components, and strongly connected components.

Strongly connected component: Each node within the component can be reached from every other node in the component by following directed links.
Weakly connected component: Each node within the component can be reached from every other node by following any direction.

Two indicators, ΔN_{WC} and ΔN_{SC}, are applied to evaluate the reduction in uncertainty and ambiguity. ΔN_{WC} refers to the AGR of the number of weakly connected components, and ΔN_{SC} refers to the AGR of the number of strongly connected components. These factors indicate the change in the topology of the citation network of the remaining topics after deleting one topic and the degree of influence of the topic on the community network.

The candidate ERTs that have a greater impact on weakly connected and strongly connected components, especially those with a weaker impact on weakly connected communities and a stronger impact on strongly connected communities, have an increasing impact on the entire knowledge network; therefore, their uncertainty also reduces. Moreover, these ERTs have a better chance of becoming a knowledge hub in the future.

3.3. Using multi-indicators in ERT detection

Simple indicators are easy to understand, but they may be highly biased, whereas complex indicators are often more precise, but they can be difficult for a general audience to interpret (Sugimoto and Larivière, 2018). In this study, we applied three groups of indicators, namely, "novelty and growth", "potential for prominent impact" (societal and economic influence), and "reduction of uncertainty and ambiguity" to detect ERTs in the field of *stem cell research*. Although novelty and growth can be estimated using several approaches, methods that include several indicators are sometimes found to be inconsistent. Therefore, a tradeoff between the different indicators is required. For example, "growth" reflects a process of accumulation and a relatively continuous trend, whereas "novelty" is a burst phenomenon that represents a discontinuous trend. Similarly, persistence and novelty are also in conflict. Persistence implies an ongoing multi-year activity, whereas novelty occurs over a relatively short period with an abrupt increase in activity. Therefore, such contradictions should be addressed according to specific situations in different domains.

Based on the analysis of the characteristics of each indicator and the correlation analysis, novelty and growth are considered the most important indicators in identifying an ERT. Therefore, these two indicators are regarded as threshold indicators. Under the premise of ensuring novelty and growth, the potential for prominent impact on society and economy and that on community network structure were considered. Hereinafter, two indicators, namely potential for prominent impact and reduction of uncertainty and ambiguity, are regarded as reference indicators (Fig. 4).

Among the threshold indicators, the growth of the number of papers can directly reflect the trend of development of a topic compared with the growth of the number of journals, funds, and authors. More notably, the recent AGR of the number of papers on a topic is a more suitable factor for analyzing the recent growth trend and novelty of the topic. The AGR of the number of journals, funds, and authors across the entire

Table 2
Retrieve strategy of patent stem cells.

No.	*Retrieve strategy*	*Remark*
1	TS = (("stem cells" OR "stem cell") NOT ("stem cellulose" or "stem. Cellular" or "cello" or "cellar" or "cellphone") NOT TS =("fuel cell" or "in-plane switching" or "Intrusion Prevention System") NOT IPC=(H or D or E or F or A23 or A46 or A24 or A47 or A63 or A62 or A44 or A45 or C07B or C07C or C07D or C07F or C07J);	conception related to stem cells
2	TS =((ESC or ASC or iPS or PGC or MSC or CSC or LSC or TSC or ADSC or HSC) near (cells OR cell)) NOT TS =("fuel cell" or "in-plane switching" or "Intrusion Prevention System") NOT IPC =(H or D or E or F or A23 or A46 or A24 or A47 or A63 or A62 or A44 or A45 or C07B or C07C or C07D or C07F or C07J);	abbreviations related stem cells research
3	TS =((("totipotent" or "pluripotent" or "multipotent" or "unipotent" or "progenitor" or "precursor") NOT ("non-pluripotent")) ADJ (cells OR cell))) NOT TS =("fuel cell" or "in-plane switching" or "Intrusion Prevention System") NOT IPC =(H or D or E or F or A23 or A46 or A24 or A47 or A63 or A62 or A44 or A45 or C07B or C07C or C07D or C07F or C07J);	synonym concept to stem cells

Note: TS denotes title and abstract.

stage measures the attention a topic receives from multiple perspectives. A topic with a large number of output contributions, a large amount of funding, and a large number of research scholars is more likely to become a key research direction in that field, naturally, with a higher forward-looking value.

For the reference indicators, the potential for a prominent impact indicator combines basic science and applied science. For a topic, the larger the proportion of papers cited by patents (or papers citing patents), the more closely the research content of the topic is integrated with science and technology, and the more likely it is to produce significant breakthrough innovations with higher socioeconomic value. Different research topics in the same field are not isolated but are related to each other. The influence of a topic on the topology of the community network formed by the citation network of other topics can reveal the importance of the topic. Topics that have a greater impact on the community network can reduce this uncertainty.

4. Case study

4.1. Data set and statistic description

An important topic in biomedical research is the study of stem cells, a type of biological cell capable of self-renewal and multidirectional differentiation. Stem cells are considered an important addition to the medical 'toolbox' and enable the development of new revolutionary disease treatments and regenerative medicine. Stem cells are a major research topic worldwide in life sciences and medical research (Hu et al., 2018). We use literature on stem cell research as a case study to demonstrate the applicability of the approach proposed in this article. The data for this case study were collected on October 20, 2018. To retrieve research articles in WoS, we use the following strategy:

TS = (("stem cells" OR "stem cell") NOT ("stem cellulose" or "stem Cellular" or "cello" or "cellar" or "cellphone")) OR TS =((ESC or ASC or iPS or PGC or MSC or CSC or LSC or TSC or ADSC or HSC) near (cells OR cell)) OR TS =((("totipotent" or "pluripotent" or "multipotent" or "unipotent" or "progenitor" or "precursor") NOT ("non-pluripotent")) ADJ (cells OR cell))) NOT TS = ("fuel cell" or "in-plane switching" or "Intrusion Prevention System")

Note: TS denotes title, abstract, author keywords and keywords plus.

To select patents from the PATSTAT database, we used the retrieving strategy listed in Table 2, separated into three groups and integrated into the final retrieving strategy.

Ultimately, 422,101 research articles and 50,556 stem cell patents were collected. The data from these articles were used in the calculation of the indicators, while the retrieved patents were only used to calculate the "potential for prominent impact" In order to avoid the influence of random fluctuations, such as expansion or reduction of the database, we use a smoothed number of publications per year. In this study, publications from the period 2004–2018 were used. This period was subdivided into eleven intervals of five consecutive years each. Adjacent intervals overlap by four years. The number of publications in each subinterval followed an exponential growth path until 2017 (Fig. 5). The decrease in numbers for the period 2014–2018 that can be observed in this figure does not necessarily mean a real decrease in the number of publications. It is merely a database artifact caused by the fact that, for 2018, not all publications were not yet available at the moment the data was collected. In addition, the fine-grained subject classification changed over time. For example, out of 4047 subject classes available at the lowest level of the Leiden classification, there are 1584 relevant

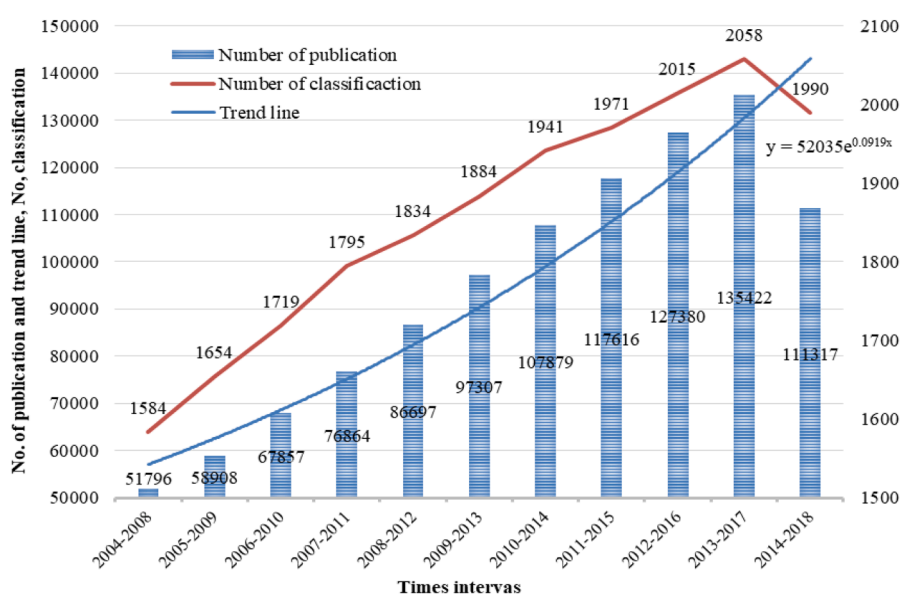

Fig. 5. Publication and classification of stem cell from 2001 to 2018.

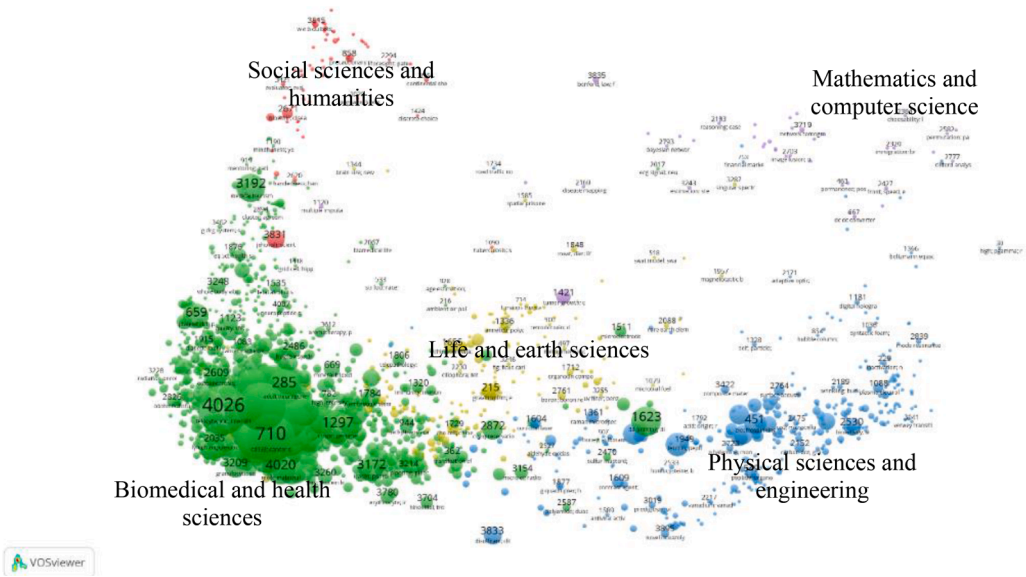

Fig. 6. Micro-level classification in stem cell from 2001 to 2018.

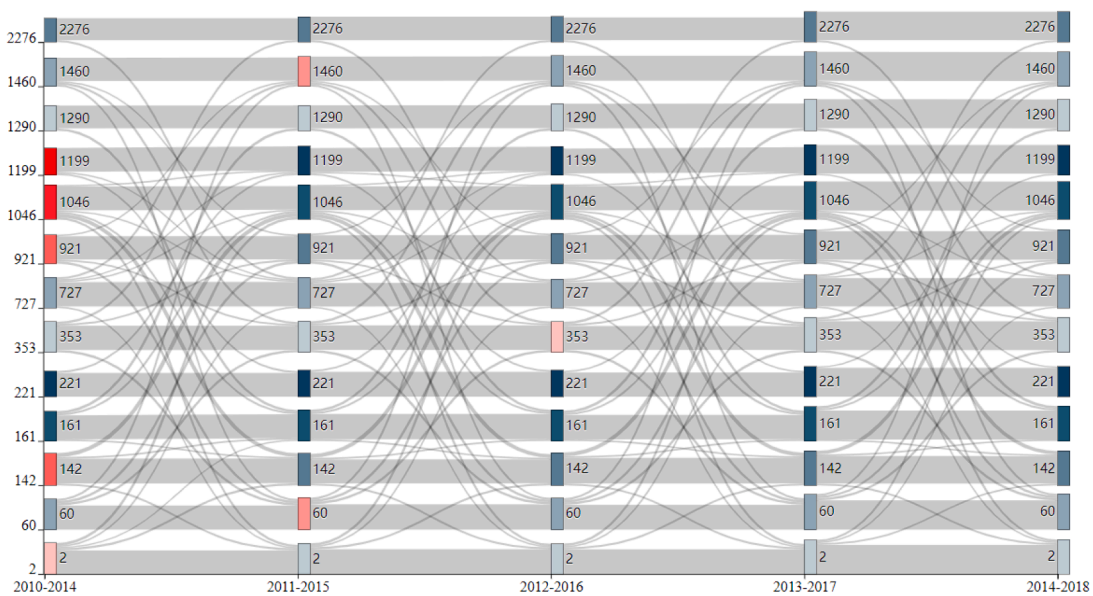

Fig. 7. The trajectories of ERTs in the stem cells field (similarity>=0.04). (https://www.informationscience.top/innovation/13.html).

classes in the period 2004–2008. The number of relevant classes increased to 2058 for the period 2013–2017 (**Error! Reference source not found.**).

4.2. Visualizing the clustering

We used VOSviewer to visualize the distribution of the classification clusters of stem cells research across science fields (Fig. 6). The size of the disk of a node for a particular classification class in this picture is proportional to the number of publications on stem cell research divided by the total number of publications for the entire WoS database in this classification class. This figure shows that the main topics are located in the "biomedical and health sciences" area. The closer two topics are placed together in the diagram, the stronger they are related. Short distances between topics indicate that they are strongly related. To find the most relevant topics in stem cell research, we obtained the topics using the Leiden-algorithm and removed topics that contained only a few stem cell publications.

For the discovery of topics, data from the WoS database was used as a source. The relevance of the generated topics is proportional to the number of publications on stem-cell research for a particular classification code divided by all publications in the entire WoS database for that classification code. To ensure that the identified ERTs are the topics that are the most directly related to stem-cell research, we analyzed for each time period the top-50 topics, based on the size of the clusters. Finally, we identified 65 unique topics in the eleven intervals. These 65 topics were further analyzed. Subjects occurring in less than three intervals were omitted from further analysis as we believe such subjects do not meet the characteristic of being an ERT in terms of persistence and coherence. Topics 1271, 376, and 1864 are among the topics excluded. In addition, we see that some topics increase slightly at first and decrease quickly later. We consider that such topics do not meet the necessary characteristics of uncertainty and ambiguity reduction. That is why we have removed them too, including topics 104, 1497, and 1576.

Table 3
Number of topics in eleven time intervals of topics 353 and 2276.

	Number of publications included in topics (N_t)										
Topic ID.	2004 −2008	2005 −2009	2006 −2010	2007 −2011	2008 −2012	2009 −2013	2010 −2014	2011 −2015	2012 −2016	2013 −2017	2014 −2018
353	–	219	340	474	616	804	1005	1154	1341	1474	1224
2276	194	225	253	306	374	451	516	618	686	807	681

Table 4
Calculation process of the growth rates of paper numbers in each interval (topics 353 and 2276).

process	formula	353		2276	
the growth rates of the number of papers	$V_T = \frac{N_{t+i} - - - N_t}{N_t}$	calculation process	result	calculation process	result
$T = 6$	$V_6 = \frac{N_7 - N_6}{N_6}$	$\frac{1005 - 804}{804}$	0.25	$\frac{516 - 451}{451}$	0.1441
$T = 7$	$V_7 = \frac{N_8 - N_7}{N_7}$	$\frac{1154 - 1005}{1005}$	0.1483	$\frac{618 - 516}{516}$	0.1977
$T = 8$	$V_8 = \frac{N_9 - N_8}{N_8}$	$\frac{1341 - 1154}{1154}$	0.1620	$\frac{686 - 618}{618}$	0.1100
$T = 9$	$V_9 = \frac{N_{10} - N_9}{N_9}$	$\frac{1474 - 1341}{1341}$	0.0992	$\frac{807 - 686}{686}$	0.1764
$T = 10$	$V_{10} = \frac{N_{11} - N_{10}}{N_{10}}$	$\frac{1224 - 1474}{1474}$	−0.1696	$\frac{681 - 807}{807}$	−0.1561

A total of eleven topics were removed during this process, as shown in Table A1 in the appendix. Of the original 65 topics, the final selection consisted of 54 that are considered potential (candidate) ERTs and were used for the multidimensional indicator identification process.

4.3. Main steps of ERTs in the stem cell field

The framework described in Section 3 was used to calculate the indicators of the five attributes for each candidate ERT. The time window for this study was 2004–2018. In this time frame, eleven 5-year subintervals were defined as noted above. The top 50 topics were selected for analysis. The specific steps are as follows:

Step 1(Novelty and growth): For each topic, the average growth rate (AGR) of the number of articles over the entire time interval was obtained by averaging the growth rates between consecutive intervals. In this way, the moment a topic becomes visible can be identified. After this identification process, topics that are more recent are given more importance. This applies in particular to the topics that appeared after 2010. Topics that disappeared after being previously visible were excluded because they apparently had less potential to become an ERT. The AGRs were calculated using Eq. (1) in Section 3.2.1.

Using eleven time periods, we can calculate ten growth rates for successive periods. The AGRs ($N_T = 10$) for the number of journals, funds, and authors for each topic were calculated for all 5-year periods. Frequency counts were used to calculate the AGRs for each topic. Topics with null values in one or more of the AGRs were excluded from the analysis because they did not fall in the top 50 of the journal, fund, or author rankings, indicating that the topic received less research attention. For example, for topic 1297, although the AGR of the number of papers ranked high, it had a null value in the AGRs of journals, funds, and authors. Thus, the topic was not influential enough and was excluded. It should be noted that this step did not necessarily exclude topics with negative growth because a negative value for a single indicator, such as journal, fund, or author AGR, was not sufficient to confirm that the topic did not have the potential to be an ERT. For example, for topic 469, the AGR of the number of journals was negative, whereas the rest of the indicators were positive, indicating that the overall growth of the topic was good. In this manner, 10 topics were removed, and 44 potential ERT technology topics were selected.

Table 6
ERTs identification process and result in stem cells study.

Operation sequence	Operational approach	Numbers of articles or topics
Data acquisition	Data acquisition	422,101 research articles
Topic discovery	Topic discovery	2000+ topics
Topic filtering	Identified ERTs topics most directly related to stem cells study	65 topics
	Removed topics only exist in a few time interval	57 topics
	Removed topics increase slightly and then decrease rapidly	54 topics
Main steps of ERTs in the stem cell field	Step 1 (Novelty and growth, $N_T = 10$)	44 topics
	Step 1 (Novelty and growth, $N_T = 5$)	26 topics
	Step 2 (Persistence and coherence)	26 topics
	Step 3 (Potential for high impact)	26 topics
	Step 4 (The reduction trend of uncertainty and ambiguity):	26 topics

Finally, to analyze the recent growth of the potential ERTs, the AGRs for the number of papers for each topic over the last five adjacent time intervals were obtained. Topics with negative growth rates for the number of publications in the last five adjacent time intervals were excluded as negative values indicate that the number of papers on this topic continues to decrease in the near future. This behavior was observed for topic 23. Although the number of journals, funds, and authors increased, the number of papers declined recently, showing that its novelty had worn off. Therefore, it was excluded from further study.

We illustrate the specific calculation process of AGRs using topics 353 and 2276. The period from 2004 to 2018 was selected as the overall

Table 5
The calculation process of AGR of paper numbers (topic 353 and 2276).

Topic ID.	Total growth rate Calculation: V	N_T	Total growth rate	Average growth rate calculating	Average growth rate calculation	AGR of paper numbers
353	$V = V_6 + V_7 + V_8 + V_9 + V_{10}$	5	0.4899	$\overline{V} = \frac{1}{5}V$	$\frac{0.4899}{5}$	0.0980
2276		5	0.4721		$\frac{0.4721}{5}$	0.0944

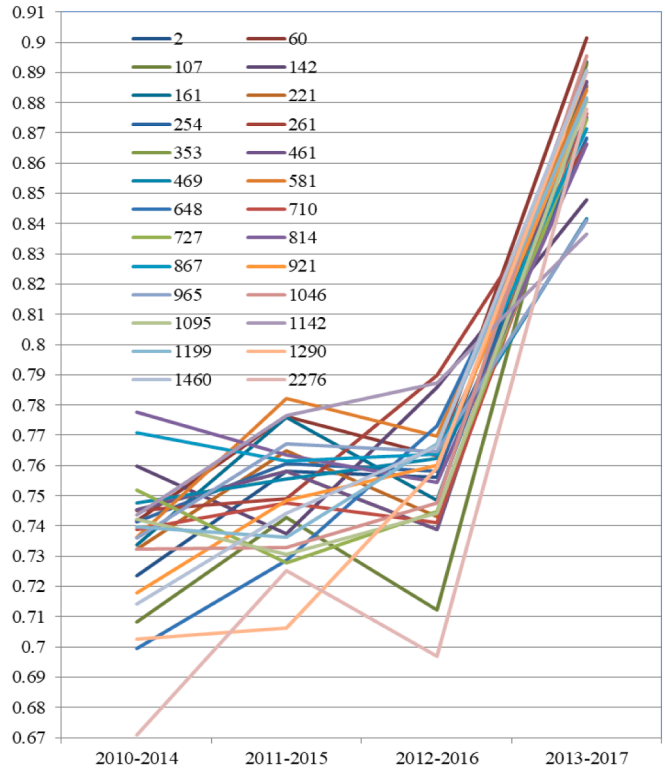

Fig. 8. Trajectories of ERTs in the stem cells field (Jaccard similarity>=0.04).

research period and was divided into eleven smaller sub-intervals, each comprising five adjacent years, as shown in Table 3.

The first calculation of the growth rate is for the time intervals 2004–2008 and 2005–2009. The second calculation uses the data for 2005–2009 and 2006–2010, and so on. To analyze the recent growth of the potential ERTs, the AGRs for the number of papers for each topic over the last five 5-year intervals were calculated. Table 4 shows the specific calculation process and results.

In the second step, the topic growth rates of topics 353 and 2276 are summed into the total growth rate. Finally, the average growth rates for topics 353 and 2276 were calculated (Table 5).

Step 2 (Persistence and coherence): Finally, 26 candidate ERTs were selected. Table A2 in the Appendix summarizes the reasons why 28 candidate topics were excluded from the final ERT list. To make the entire empirical analysis process clearer, Table 6 further summarizes the process list of ERT topics from the data acquisition to finally identify the ERTs of this study.

Error! Reference source not found. illustrates the evolution paths over the last five 5-year intervals for ten representative ERTs in stem cell research with Jaccard coefficients ≥0.04. This diagram was constructed using D3.js software (Bostock et al., 2011). The horizontal axis is the time span, and the vertical axis is the topic ID. In Figure 7., higher similarities are indicated by thicker lines, and the red colored topics have a higher potential to be an ERT. Each topic follows its main trajectory, indicating that ERTs will continue to exist. Topics that intersect show the coherent nature of ERTs. ERTs showing a continuous upward trend are expected to produce future socioeconomic value.

Fig. 8 shows all 26 candidate ERTs for stem-cell research. The coherence degrees (Jaccard similarity) of 26 candidate ERTs show an increasing trend, especially during the last time interval. Specifically, for all 26 candidate ERTs, the coherence degree of eight topics increased over all four intervals, 17 topics increased over three intervals, and one topic increased over two intervals. Table A3 in the Appendix provides details about the degree of coherence.

Step 3 (Potential for high impact): For each topic, two indicators, U_{BA} and U_{AB}, were calculated to evaluate the potential of high impact (see 3.2.2.). Higher values of both U_{BA} and U_{AB} indicate that a candidate ERT has more potential to have a high socioeconomic impact. We calculated the linkage between basic science and applied science based on citation relationships. The data of citations between papers and patents were acquired from the WoS and PATSTAT (EPO Worldwide Patent Statistical Database) databases of CWTS (in-house version) (Winnink, 2018).

Step 4 (*The reduction trend* of uncertainty and ambiguity): First, the citation network for all topics in the eleven intervals was constructed. Figure 9 shows the complete citation network from 2004 to 2008.

First, for each time interval, the number of weakly and strongly connected components in the entire citation network was obtained by removing candidate ERTs one at a time. Then, the absolute change of the weakly and strongly communities are obtained for all candidate ERTs, and the time series of the differences between the weakly and strongly

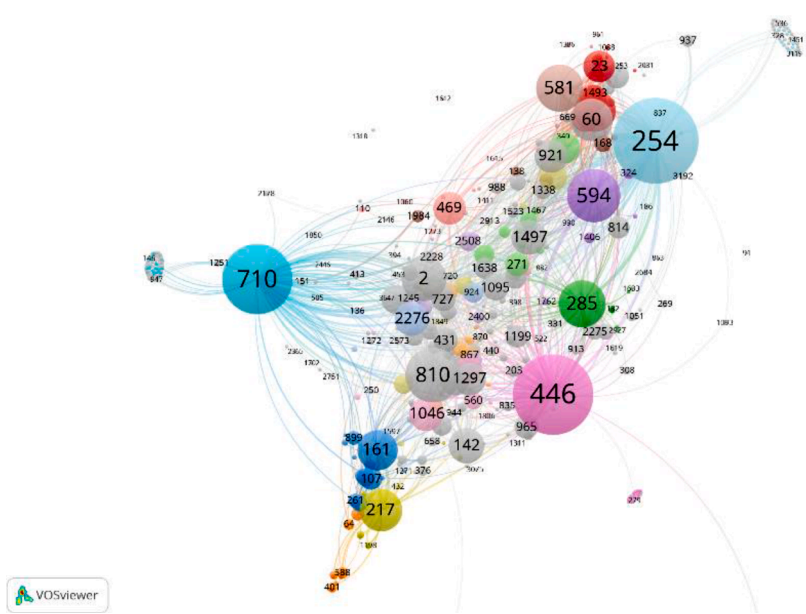

Fig. 9. Topic citation network (2004–2008).

Table 7
Uncertainty and ambiguity of candidate ERTs in stem cells field (partial).

No.	Topic ID.	ΔN_{SC}	ΔN_{WC}
1	353	1.0833	−0.1847
2	2276	0.4375	−0.0083
3	1460	−0.25	−0.5
4	142	35.3759	0.0833
5	60	−0.0833	1
6	921	0.125	0.2083
7	727	—	−0.5
8	1046	0	—
9	161	0.2125	0.125
10	221	−0.25	0

Note: "—"indicates that it is not applicable because the denominator for calculating the average rate of change is 0. For example, the change in the number of weak ties in the last five stages for topic 1046 is 0, 0, 0, 0, 1.

connected for each candidate ERT was calculated. Finally, ΔN_{WC} and ΔN_{SC} are calculated for each candidate ERT for all eleven time intervals. The values of the ten topics are listed in Table 7, and the others in Table A4 in the Appendix.

4.4. ERTs in the stem cell field

In the end, this study identified 26 ERTs in the stem cell research field. The features of the ten topics are listed in Table. 8, and the features for the other topics can be found in Table A4 in the Appendix.

4.5. Correlation analysis of ert identification indicators

A correlation analysis for all eight indicators (AGR of paper numbers, AGR of journal numbers, AGR of fund numbers, AGR of author numbers, U_{BA}, U_{AB}, ΔN_{WC}, and ΔN_{SC}) was performed. In this study, the higher the value of all eight indicators, the more likely it is that a topic is a potential ERT. The correlation between the indicators may provide some insights to identify ERT by using multi-indicators. Based on the ERT identification indicators in the field of stem cell research, we calculated the correlation between these indicators.

Table.9 presents the results of Pearson correlation analysis for the eight secondary indicators by SPSS 20.0, where only the change in the number of funds and U_{BA} have a significant positive correlation at the 0.01 level. The other indicators show no significant correlations, indicating that most indicators are independent, especially the first-level indicators that require cross-referencing.

4.6. Evaluation analysis

To verify the validity of the method, we compared the prediction results with the predictions of domain experts in stem cell research and also compared the results of the empirical analysis with qualitative evidence in the field of stem cells, including authoritative review of prospective articles, annual scientific breakthroughs, etc. These considerations, as presented in Table 10, can serve as evidence of the effectiveness of the proposed method.

Applying the Leiden algorithm (Traag et al., 2018) to the entire WoS database, the generated detailed classification could assist in the discovery of more accurate topics. Nonetheless, the classification still lacks specificity. Therefore, to further gain an in-depth understanding of the content of each ERT, we conducted an additional topic analysis of the text covering the resulting study area. We then re-labeled them more suitably. Table A5 in the Appendix depicts the new topic labels for the ERTs of the stem cell field.

4.7. ERTs patterns and R&D layout strategy

The 26 topics we identified are ERTs with high novelty and growth in the stem cell field. However, several obvious differences can be observed in the performance of the reference indicators. At the same time, the socioeconomic influence and the future development uncertainty in the reference indicators are related to R&D layout strategies and development patterns. Therefore, we further combined the threshold and reference indicators to divide the 26 ERTs into different ERT patterns, and different development strategies are suggested for the R&D layout of

Table 8
ERTs in stem cells field (partial).

		(i) Novelty(ii) Growth				*(iii)Potential prominent impact*		*(iv)Reduction to uncertainty and ambiguity*	
No.	*Topic ID.*	*AGR of paper numbers*	*AGR of journal numbers*	*AGR of funding numbers*	*AGR of authors numbers*	U_{BA}	U_{AB}	ΔN_{WC}	ΔN_{SC}
1	353	0.0980	0.1999	0.2367	0.1088	0.0745	0.0082	−0.1847	1.0833
2	2276	0.0944	0.1121	0.5297	0.0694	0.1061	0.0051	−0.0083	0.4375
3	1460	0.0669	0.1660	0.6678	0.0624	0.111	0.0202	−0.5	−0.25
4	142	0.0598	0.1987	0.4830	0.1166	0.1191	0.0014	0.0833	35.3759
5	60	0.058	0.0975	0.5135	0.0636	0.113	0.0086	1	−0.0833
6	921	0.0551	0.0748	0.5400	0.0669	0.0777	0.0015	0.2083	0.125
7	727	0.0526	0.0883	0.4257	0.0807	0.1539	0.0104	−0.5	—
8	1046	0.051	0.1400	0.5461	0.1455	0.0972	0.0091	—	0
9	161	0.0422	0.0358	0.5476	0.0615	0.0937	0.0015	0.125	0.2125
10	221	0.0419	0.0539	0.4183	0.0364	0.1423	0.0018	0	−0.25

Table 9
Correlation Analysis of ERT Identification Indicators.

	AGR of paper numbers	AGR of journal numbers	AGR of fund numbers	AGR of author numbers	U_{BA}	U_{AB}	ΔN_{WC}	ΔN_{SC}
AGR of topic numbers	1	0.26	0.136	−0.04	−0.115	−0.035	0.046	−0.049
AGR of journal numbers	0.26	1	−0.805**	0.339	−0.906**	0.313	−0.33	−0.085
AGR of funding numbers	0.136	−0.805**	1	−0.328	.874**	−0.123	0.189	0.183
AGR of authors numbers	−0.04	0.339	−0.328	1	−0.326	−0.095	0.109	0.056
U_{BA}	−0.115	−0.906**	.874**	−0.326	1	−0.131	0.175	0.277
U_{AB}	−0.035	0.313	−0.123	−0.095	−0.131	1	−0.29	0.217
ΔN_{WC}	0.046	−0.33	0.189	0.109	0.175	−0.29	1	−0.029
ΔN_{SC}	−0.049	−0.085	0.183	0.056	0.277	0.217	−0.029	1

*. Significantly correlated at the 0.05 level (both sides).
**. Significantly correlated at the 0.01 level (both sides).

Table 10
Major progress of ERTs in stem cells field (partial).

Topic id	*Emergent evidences for ERTs in stem cells research*
353	September 2018, Sullivan et al. (2018) published research results in the journal *Cell* and found that the extracellular matrix can control the movement of cells in the body by regulating the cell's glucose metabolism. This study is of great significance to reveal the metastasis of cancer cells.
814	According to the records of ClinicalTrials.gov, there are currently 50 international clinical trials of stem cells for the treatment of spinal cord injury, involving bone marrow stem cells, mesenchymal stem cells, autologous stem cell transplantation and other research directions (ClinicalTrials.gov, 2019). In May 2018, E. Curtis et al. (2018) published research results on the use of neural stem cells to treat spinal cord injury, and 3 of 4 subjects had a significant improvement in their condition. In August 2018, K. Hiromi et al. (2018) published research results in the subjournal *Nature*, and created spinal cord-derived neural stem cells through human pluripotent stem cells (hPSCs), which is expected to be used in the clinical treatment of spinal cord injury.
1460	In September 2010, professors John Fisher and Eileen Ingham of the University of Leeds in the United Kingdom made naked scaffolds from passed donors or parts of animals' bodies, and successfully used patient stem cells to make new body parts. In 2016, J. P. Guyette et al. (2016) used human induced pluripotent stem cells as seed cells to inoculate heart decellularized scaffolds from organ donors to construct tissue engineered hearts. In 2010, Shay Soker, the director and professor of Wake Forest University School of Medicine with his team members, used human stem cells to make miniature human livers. This technology retains tissue or cell-specific ECM protein components and microstructures, strips materials that trigger rejection reactions, and is widely used (Bioon, 2018).
254	Until March 2019, a total of 38 stem cell clinical research projects in China have been filed, of which 29 clinical projects have used mesenchymal stem cells (Youzhunet, 2019).
648	Using patient autologous cells or amniotic cells, limbal stem cell transplantation can successfully repair functional corneal epithelium. Autologous limbal stem cell expansion cultured cell therapy has been approved for use in the European Union (iyiou, 2019). This technology has broad application scenarios.
161	More and more patients are receiving allogeneic hematopoietic stem cell transplantation to treat acute myeloid leukemia.
261	Sequential autologous hematopoietic stem cell transplantation with high-dose chemotherapy has become the standard treatment option for relapsed / refractory Hodgkin's lymphoma (Wang, 2016).

Table 11
Different ERTs patterns and R&D layout pattern.

Patterns	Topic	Scientific layout pattern
NG-PI-UR	1046	Focus on and increase in investment and policy support, step up scientific and technological layout, and occupy the commanding heights of science and technology
NG-PI	1460	Prioritize layout, focus on investment, monitor uncertainty in future development
NG-UR	142	Prioritize layout, guide the transformation of scientific and technological achievements, promote the efficient use of scientific and technological achievements, and bring out the societal and economic benefits
PI-UR	581; 469	Appropriate layout, promote cooperation between industry, universities and research institutes, monitor novelty and growth rate, and explore the topic's forward-looking value
NG	2; 2276	Appropriate layout, accelerate the transformation of achievements, and monitor the uncertainty of future development
PI	254; 727; 1142; 107	Appropriate layout, focus on novelty and growth rate
UR	1290; 60; 353; 461; 814; 221; 261	Encourage cooperation between industry, universities and research institutes, explore the research value of the topic and its societal and economic benefits

Note: three topics without obvious topics patterns.

the resulting ERT.

Topics with a high novelty and growth rate reflect the continuing interest of the academic community in this ERT and are part of the frontier of scientific research. Therefore, it is necessary to create and follow an R&D layout in a timely manner. The ERTs with high technology interaction reflect a good transfer of knowledge between basic research and applied research. This topic pattern may have high socioeconomic value, rather than just remaining in the stage of basic research. Therefore, these topics should inform research prioritization for R&D planning in regional economic development. Finally, ERTs with low uncertainty can be guided to encourage more cooperation between industry, university, and research institutes, so as to accelerate the development of ERTs to realize their socioeconomic value through collaborative innovation.

By combining three primary indicators, (i) novelty–growth (NG), (ii) potential for prominent impact (PI), and (iii) reduction of uncertainty and ambiguity (UR), the 26 ERTs can be divided into the following seven different types (Table. 11):

(1) High novelty and growth rate, higher potential socioeconomic impact, and less future development uncertainty (NG-PI-UR);
(2) High novelty and growth rate and higher potential socioeconomic impact (NG-PI);
(3) High novelty and growth rate and less future development uncertainty (NG-UR);
(4) High potential socioeconomic impact and less future development uncertainty (PI-UR)
(5) High novelty and growth rate (NG);
(6) Higher potential socioeconomic impact (PI);
(7) Less future development uncertainty (UR).

Topic 1046 is the only ERT with the NG-PI-UR pattern. Not only the novelty and growth indicator scored high, but also the proportion of publications in which patents were mentioned and the changes in the number of research communities. This indicates that the topic is not intensively studied, but also that basic research and applied research closely interact. This indicates that it is likely to have high research value and socioeconomic benefits. At the same time, this topic is also the core and key structure in the knowledge network, which may require critical focus and R&D input; it is also the priority of the technology to develop the R&D layout.

Topic 1460 is the sole ERT with an NG-PI pattern. It ranked first for the AGRs of the number of papers and funds. The proportion of papers citing patents indicates that topic 1460 is not only novel, but also highly forward-thinking. At the same time, the applied research on the topic has an obvious effect in the promotion of basic research, which may have a significant impact on society and the economy. However, as its future development has a higher uncertainty, this topic needs to be monitored in future development trends when planning R&D input.

Topic 142 is taken here as an example NG-UR pattern. It ranked highest for the novelty and growth indicator, and the impact on the number of strongly connected components in the community network structure is much higher than that of other ERTs. The topic is in an important position in the community network and has a high degree of structural relevance to other topics, thus reducing the uncertainty of its future development. It may be a knowledge hub in the research of stem cells, and consequently needs to be focused and guided in the R&D layout. At the same time because the potential socioeconomic impact of the topic is not high, it is necessary to promote efficient transformation of scientific and technological achievements to industrialization and promote the realization of its economic and societal effects.

PI-UR pattern ERTs, such as topic 469, have a low novelty and growth indicator rank, but involve the fundamental transfer of basic research to applied research. Since the proportion of papers cited by patents and the impact on the number of strongly connected components in the community, network structure indicators are ranked high and are

in a key position in the community network, reducing the uncertainty of future development. These topics need appropriate R&D layout, guidance, and cooperation to promote knowledge transfer between industry, universities, and research institutes. More exploration value of the research needs to be done.

Topic 2 is an example of an NG pattern ERT. It had a high degree of novelty, but the potential socioeconomic impact and the impact on the community network structure are not very obvious. For NG pattern topics, it is necessary to monitor their future development. To reduce the uncertainty of such a topic, an R&D layout strategy that accelerates the transfer of research achievements is beneficial.

Topic 107 is taken as an example of the PI pattern. The AGR of the number of papers, the number of funds, the number of journals, and the number of authors ranked low, but the proportion of papers cited by patents ranked first, indicating that the topic does not have the highest novelty ranking, but the basic research is well-suited for the development of applied research. High-level basic research achievements are of great significance for promoting high-tech development and technological innovation. Therefore, these topics need close attention to encourage collaboration between industry, universities, and research institutes, and to further explore their research value and socioeconomic benefits.

Topic 1290 is an example of the UR pattern. It ranked first among the 26 topics for the degree of change in both the number of communities and weakly connected components. It ranked second for the strongly connected component indicator, indicating that this topic holds an important role in the knowledge network structure, and is therefore a key research topic in the field. It needs to be paid close attention, with the intention to discover potential cooperation partners between industry, universities, and research institutes. It would be beneficial to promote the interaction between basic research and applied research, timely R&D layout, and exploration of the research value and socioeconomic benefits of the topic as early as possible.

5. Discussion and conclusions

5.1. Conclusions

This study introduces a multidimensional and practical bibliometric methodology for the detection of ERTs. The proposed framework facilitates the detection of ERTs at a fine level of topic granularity. To instantiate this approach, an in-depth analysis was conducted in the field of stem-cell research. In this study, we focused on the potential high impact and the uncertainty or ambiguity of an ERT. Therefore, our approach could provide more comprehensive information about ERTs. Additionally, this approach was found to be flexible because it could be customized according to the concrete features of specific domains. The threshold indicators included AGRs, the number of candidate ERTs, and the Jaccard similarity for the evolution map. We further divided the identified ERTs into seven patterns and provided a guide for different R&D layout strategies for the different ERT patterns. Therefore, the method outlined herein is expected to provide significant support for science and technology policy development.

We compared the empirical results with other types of evidence from the stem cell field. The evidence included information from the research progress of different ERT technologies published in professional journals, newsletters, and printed domain expert opinions. By combining this information, we demonstrated that our method could effectively identify ERT fields. Additionally, our discovered ERTs were supported by recent major news reports and associated papers. These facts indicate that the methodology can be used by decision makers for the evaluation of scientific research.

5.2. Main contributions

The main contributions of this study are:

1 This study operationalized multidimensional scientometric indicators for ERTs, including prominent impact, development uncertainty, ambiguity, radical novelty, relatively rapid growth, and coherence. Five attributes were measured simultaneously to detect ERTs, which, in our opinion, was the first time to do so.
2 Although many studies have been conducted to identify ERTs, the concept of ERT remains ambiguous. We used qualitative and quantitative methods to make the connotations and extensions of the ERT and related concepts less ambiguous.
3 The analysis of extensive experimental stem cell results confirmed the feasibility and effectiveness of our method for the use of multidimensional scientometric indicators for the detection of ERTs.
4 The study identified different R&D layout strategies for each indicator's development pattern.

5.3. Limitation

Several limitations and deficiencies were observed in this study. First, although the timing of a research topic and the changes in the number of papers, journals, funds, and authors involved in the topic can characterize its novelty and growth, the measurement indicators may not form an unbiased mapping because ERTs can exist and spread through implicit knowledge dissemination channels other than scientific publications. In this case, it is quite possible that related publications might not be publicly accessible.

Second, the mutual citation of scientific papers and patent documents reflects the interaction between basic research and applied research, forming an assessment of the impact on the economy and the societal aspects of scientific and technological development. However, this does not necessarily present a positive causal relationship in which a topic reflects a good interaction between basic research and applied research. Whether the topic will indeed become an ERT that promotes economic and societal development is conditioned by various factors not discussed here, including restrictions to the current level of technology and policy implications.

Third, the central influence of the topic in the network structure used to detect the uncertainty of ERTs was derived only from the topology of the knowledge node in the network. However, the real evaluation of uncertainty reduction is far more complicated than topology analysis, which is usually affected by more factors (e.g., policies, ethics, and opportunities).

As noted by Bolger and Wright (2017), there are still no repetitive phenomena or events that can be used as a reference point for improving the accuracy of detection and forecasting. Therefore, it is almost impossible to compare directly detected and forecasted emerging technologies with realized outcomes. This point is also applicable to our case. Although Apreda et al. (2019) re-examined the findings of a technology foresight exercise on the medical-device industry with the realized technological performance five years later, this problem was still very difficult to address. Huang et al. (2019) compared recent technology activities with previous forecasts to reveal the influencing factors that led to differences between past predictions and actual performances. They found that data quality, forecasting methods, complexity, and dynamism of future-oriented analyses could cause differences between preceding predictions and actual performance.

5.4. Future work

In the near future, our framework will be augmented with new and unconventional methods and various information resources to identify transformative ERTs. Similar to Huang et al. (2019), we plan to conduct a revisiting assessment research by the end of 2023, which is five years after 2018, when there will be a complete and independent data range, with five adjacent years from 2019 to 2023. Therefore, we can test the accuracy of multidimensional indicators and further reveal the influencing factors that lead to differences between past predictions and

actual performance, and accordingly make further improvements to our multidimensional indicators.

This study only considered scholarly publications and patent documents. The utilization of comprehensive multi-source information could result in a more objective result, as different types of data resources reflect various technological perspectives (Xu et al., 2017).A collection of a variety of data sources, such as conference information, scientific projects, reports, and industrial economic data, should be included in the scope of scientometric analysis. For example, industrial economic data helps to grasp technology market information.

As noted by Rotolo et al. (2015), scientometric indicators cannot comprehensively evaluate all ERT attributes. In the near future, the potential high impact of ERTs will emphasize the economic viewpoint. Altmetrics (e.g., social media, newspapers, and magazines), policy reports, and scholarly publications in the social sciences that mention ERTs may provide a feasible solution. Unlike the mature methods of radical novelty, relatively fast growth, and coherence, research on potential high impact, uncertainty, and ambiguity reduction is still at the initial exploratory stages. Hence, we have not yet shown its true advantages. These two aspects require more attention in future research. The method developed in this study is only a tentative and feasible trial. To evaluate the uncertainty or ambiguity of an ERT, indicators and methods other than scientometrics such as Kalman filters (Sheen and Wang, 2017) and entropy-based methods (López–Menéndez and Pérez-Suárez, 2019; López–Menéndez and Pérez-Suárez, 2020) should be utilized in economic uncertainty prediction.

It is well known that early identification of ERTs has important strategic significance for promoting science and technology development. In this aspect, our framework still has a large room for improvement. An ERT usually exists in the form of a weak signal during the initial budding stage. Many weak signals appear as noise, but they are still indicators of significant future events or trends when combined with other types of information (Day and Schoemaker, 2006; Schoemaker et al., 2013). Unlike strong signals, such as highly cited documents and high-frequency words, weak signals contain more diverse information. They can reflect the knowledge structure of different disciplines (Wei et al., 2017). Therefore, weak-signal tracing and analysis can help us recognize ERTs as early as possible. Our future research will focus on capturing effectively weak signals.

Moreover, most current ERT identification methods pay more attention to progressive innovation, and identification methods for transformative topics (TTs) are still under investigation. Although progressive innovation research usually adheres to a relatively stable development path, a development pattern for TTs is more difficult to obtain due to its disruptive and nonlinear nature. Especially during the early days of TTs research, obvious signs were very difficult to detect. Therefore, in the future, TTs identification studies will use a different methodology for progressive ERT identification using some less-conventional methods.

Credit author statement

Haiyun Xu: (xuhy@clas.ac.cn, corresponding author): - Investigation, - Methodology, - Validation, - Visualization, Writing - original draft, - review & editing.

Jos Winnink: (winninkjj@cwts.leidenuniv.nl): - Methodology, - Validation, - review & editing.

Zenghui Yue: (yzh66123@126.com): - Formal analysis, - original draft.

Huiling Zhang: (lingkuailewuxian@sina.com): - Formal analysis, - review & editing.

Hongshen Pang: (phs@szu.edu.cn): - Resources, - review & editing.

Acknowledgments

We appreciate the constructive instruction and suggestions from Professor Ludo Waltman at the centre for Science and Technology Studies (CWTS), Leiden University, and Professor Shuo Xu at the College of Economics and Management, Beijing University of Technology. This article is an outcome of Tai Shan Scholar Foundation, the project "Study on the Recognition Method of Innovative Evolving Trajectory based on Topic Correlation Analysis of Science and Technology" (No. 71704170) and the project "Study on Regularity and Dynamics of Knowledge Diffusion among Scientific Disciplines" (No. 71704063), both supported by the National Natural Science Foundation of China, the China Postdoctoral Science Foundation funded project (2016M590124), and the National Social Science Fund of China (No. 19BTQ074).

Appendix

Table. A1, Table. A2, Table. A3, Table. A4, Table. A5

Table A1
Reasons for removed topics.

Reasons category	Topic	Number of publications included in the topics										
Reasons for removed topics	ID.	2004 −2008	2005 −2009	2006 −2010	2007 −2011	2008 −2012	2009 −2013	2010 −2014	2011 −2015	2012 −2016	2013 −2017	2014 −2018
Not meeting the persistence and coherence of ERTs	376	193	210	–	–	–	–	–	–	–	–	–
	580	180	–	–	–	–	–	–	–	–	–	–
	1038	180	–	–	–	–	–	–	–	–	–	–
	1271	200	–	–	–	–	–	–	–	–	–	–
	1338	–	–	–	–	–	–	–	–	–	–	429
	1397	183	–	–	–	–	–	–	–	–	–	–
	1864	–	–	250	–	325	372	–	–	–	–	–
	2228	197	215	–	–	–	–	–	–	–	–	–
Not meeting the feature of uncertainty and ambiguity reduction of ERTs	104	–	–	–	–	–	–	–	–	501	564	472
	1497	–	–	–	–	–	–	–	–	550	800	735
	1576	–	–	–	–	–	–	–	–	503	554	463

Note: "-" represents a topic that has not entered the top 50 topics in this period, so the number of documents is empty.

Table A2
Reasons for topics excluded from the final ERTs list.

No.	Topic ID.	AGR of topic numbers	AGR of journal numbers	AGR of funding numbers	AGR of authors numbers
1	1297	0.158	–	–	–
2	916	0.146	–	–	–
3	1306	0.12	0.103	–	–
4	659	0.074	0.079	0.276	–
5	686	0.046	0.062	–	–
6	850	0	0.089	–	–
7	1267	0	0.076	–	–
8	2275	0	0.013	–	–
9	467	−0.001	0.069	0.504	0.027
10	1265	−0.002	0.063	0.443	−0.115
11	217	−0.008	0.038	0.517	−0.009
12	810	−0.012	0.048	0.454	0.02
13	588	−0.012	0.027	0.444	0.03
14	446	−0.015	0.12	0.515	0.07
15	271	−0.016	0.037	0.371	−0.017
16	594	−0.017	0.063	0.456	0.004
17	401	−0.035	0.03	0.629	0.032
18	285	−0.042	0.045	0.387	−0.014
19	899	−0.05	0.034	0.397	−0.104
20	1638	−0.076	0.044	0.523	−0.042
21	1623	−0.109	0.085	0.151	–
22	23	−0.133	0.05	0.231	0.166
23	2508	−0.151	0.104	0.739	−0.036
24	451	−0.161	–	0.126	–
25	215	−0.1913	0.079	0.5	0.115
26	168	−0.197	0.141	0.797	−0.198
27	203	−0.207	0.03	0.821	−0.033
28	875	−0.23	0.006	0.621	−0.092

Table A3
Coherence degree of ERTs in stem cells field in each intervals.

Topic id	Topic id	time interval	time interval	coherence degree
2	2	2010 - 2014	2011 - 2015	0.7235
2	2	2011 - 2015	2012 - 2016	0.7581
2	2	2012 - 2016	2013 - 2017	0.7561
2	2	2013 - 2017	2014 - 2018	0.8934
60	60	2010 - 2014	2011 - 2015	0.7416
60	60	2011 - 2015	2012 - 2016	0.7762
60	60	2012 - 2016	2013 - 2017	0.7632
60	60	2013 - 2017	2014 - 2018	0.9013
107	107	2010 - 2014	2011 - 2015	0.7083
107	107	2011 - 2015	2012 - 2016	0.7429
107	107	2012 - 2016	2013 - 2017	0.7122
107	107	2013 - 2017	2014 - 2018	0.8928
142	142	2010 - 2014	2011 - 2015	0.7599
142	142	2011 - 2015	2012 - 2016	0.7373
142	142	2012 - 2016	2013 - 2017	0.7858
142	142	2013 - 2017	2014 - 2018	0.848
161	161	2010 - 2014	2011 - 2015	0.7338
161	161	2011 - 2015	2012 - 2016	0.7759
161	161	2012 - 2016	2013 - 2017	0.7484
161	161	2013 - 2017	2014 - 2018	0.8855
221	221	2010 - 2014	2011 - 2015	0.7324
221	221	2011 - 2015	2012 - 2016	0.7649
221	221	2012 - 2016	2013 - 2017	0.7431
221	221	2013 - 2017	2014 - 2018	0.8855
254	254	2010 - 2014	2011 - 2015	0.7415
254	254	2011 - 2015	2012 - 2016	0.7606
254	254	2012 - 2016	2013 - 2017	0.7581
254	254	2013 - 2017	2014 - 2018	0.8684
261	261	2010 - 2014	2011 - 2015	0.7452
261	261	2011 - 2015	2012 - 2016	0.7491
261	261	2012 - 2016	2013 - 2017	0.7899
261	261	2013 - 2017	2014 - 2018	0.8664
353	353	2010 - 2014	2011 - 2015	0.7454
353	353	2011 - 2015	2012 - 2016	0.758
353	353	2012 - 2016	2013 - 2017	0.7389

Table A3 (*continued*)

Topic id	Topic id	time interval	time interval	coherence degree
353	353	2013 - 2017	2014 - 2018	0.887
461	461	2010 - 2014	2011 - 2015	0.754
461	461	2011 - 2015	2012 - 2016	0.7639
461	461	2012 - 2016	2013 - 2017	0.7601
461	461	2013 - 2017	2014 - 2018	0.8792
469	469	2010 - 2014	2011 - 2015	0.7477
469	469	2011 - 2015	2012 - 2016	0.7555
469	469	2012 - 2016	2013 - 2017	0.7623
469	469	2013 - 2017	2014 - 2018	0.8417
581	581	2010 - 2014	2011 - 2015	0.7362
581	581	2011 - 2015	2012 - 2016	0.7821
581	581	2012 - 2016	2013 - 2017	0.7698
581	581	2013 - 2017	2014 - 2018	0.8841
648	648	2010 - 2014	2011 - 2015	0.6994
648	648	2011 - 2015	2012 - 2016	0.7286
648	648	2012 - 2016	2013 - 2017	0.773
648	648	2013 - 2017	2014 - 2018	0.875
710	710	2010 - 2014	2011 - 2015	0.7387
710	710	2011 - 2015	2012 - 2016	0.7476
710	710	2012 - 2016	2013 - 2017	0.741
710	710	2013 - 2017	2014 - 2018	0.8764
727	727	2010 - 2014	2011 - 2015	0.7519
727	727	2011 - 2015	2012 - 2016	0.7277
727	727	2012 - 2016	2013 - 2017	0.7447
727	727	2013 - 2017	2014 - 2018	0.8746
814	814	2010 - 2014	2011 - 2015	0.7776
814	814	2011 - 2015	2012 - 2016	0.7634
814	814	2012 - 2016	2013 - 2017	0.7543
814	814	2013 - 2017	2014 - 2018	0.8663
867	867	2010 - 2014	2011 - 2015	0.7708
867	867	2011 - 2015	2012 - 2016	0.7614
867	867	2012 - 2016	2013 - 2017	0.7637
867	867	2013 - 2017	2014 - 2018	0.8714
921	921	2010 - 2014	2011 - 2015	0.718
921	921	2011 - 2015	2012 - 2016	0.7484
921	921	2012 - 2016	2013 - 2017	0.7601
921	921	2013 - 2017	2014 - 2018	0.8811
965	965	2010 - 2014	2011 - 2015	0.7359
965	965	2011 - 2015	2012 - 2016	0.7671
965	965	2012 - 2016	2013 - 2017	0.7647
965	965	2013 - 2017	2014 - 2018	0.8412
1046	1046	2010 - 2014	2011 - 2015	0.7322
1046	1046	2011 - 2015	2012 - 2016	0.733
1046	1046	2012 - 2016	2013 - 2017	0.7476
1046	1046	2013 - 2017	2014 - 2018	0.8955
1095	1095	2010 - 2014	2011 - 2015	0.7422
1095	1095	2011 - 2015	2012 - 2016	0.7305
1095	1095	2012 - 2016	2013 - 2017	0.7443
1095	1095	2013 - 2017	2014 - 2018	0.8807
1142	1142	2010 - 2014	2011 - 2015	0.7437
1142	1142	2011 - 2015	2012 - 2016	0.7766
1142	1142	2012 - 2016	2013 - 2017	0.7873
1142	1142	2013 - 2017	2014 - 2018	0.8365
1199	1199	2010 - 2014	2011 - 2015	0.7398
1199	1199	2011 - 2015	2012 - 2016	0.7362
1199	1199	2012 - 2016	2013 - 2017	0.7671
1199	1199	2013 - 2017	2014 - 2018	0.8816
1290	1290	2010 - 2014	2011 - 2015	0.7027
1290	1290	2011 - 2015	2012 - 2016	0.7062
1290	1290	2012 - 2016	2013 - 2017	0.7593
1290	1290	2013 - 2017	2014 - 2018	0.89
1460	1460	2010 - 2014	2011 - 2015	0.7141
1460	1460	2011 - 2015	2012 - 2016	0.7441
1460	1460	2012 - 2016	2013 - 2017	0.7662
1460	1460	2013 - 2017	2014 - 2018	0.8906
2276	2276	2010 - 2014	2011 - 2015	0.6709
2276	2276	2011 - 2015	2012 - 2016	0.7251
2276	2276	2012 - 2016	2013 - 2017	0.6968
2276	2276	2013 - 2017	2014 - 2018	0.8779

Table A4
Indicators of ERTs in stem cells field.

No.	Topic ID.	*(i) Novelty,(ii) Growth* AGR of topic numbers	AGR of journal numbers	AGR of funding numbers	AGR of authors numbers	*(iii)Potential prominent impact* U_{BA}	U_{AB}	ΔN_C	ΔN_{WC}
2	0.0763	0.4144	0.6387	0.3309	0.1208	0.0016	0	−0.2083	−0.125
1290	0.0398	0.0507	0.6679	0.0669	0.0629	0.0005	7.3333	−1	9.5
1199	0.0271	0.1241	0.4037	0.0343	0.0824	0.0017	−0.25	0	–
710	0.0232	0.1162	0.516	0.1908	0.1046	0.0039	0.0118	−0.0044	−0.0512
814	0.0209	0.0443	0.4722	0.0595	0.0957	0.0045	0.3333	−0.25	−1
254	0.0188	0.0788	0.4956	0.1233	0.1716	0.0052	0.1406	0.1147	0.0162
581	0.0164	0.0611	0.5123	0.0575	0.11	0.0111	1.1917	0.4417	−0.0869
1142	0.0156	0.0507	0.502	0.0659	0.164	0.0032	−0.3542	−0.425	−0.6667
461	0.0146	0.037	0.5276	0.0547	0.08	0.0013	−1	−1	−0.5
261	0.0136	0.0008	0.4096	0.043	0.0396	0.0004	−1	−0.5556	0
867	0.0111	0.0724	0.3868	0.0265	0.0291	0	−0.0556	0.0625	0.1111
648	0.0077	0.0316	0.436	0.0466	0.1193	0.0065	−0.1667	0	−0.5
1095	0.0072	0.047	0.5289	0.0562	0.1361	0.0043	0	−0.125	−0.125
107	0.004	0	0.4249	0.0039	0.2617	0.0089	0	–	–
965	0.0037	0.1131	0.446	0.0378	0.1075	0.0057	−0.1111	0.5	0.2917
469	0.0023	−0.0174	0.402	0.0338	0.1609	0.0041	0.375	0.1458	1.0833

Note: "—" indicates it does not exist because the denominator in the formula for calculating the average rate of change cannot be calculated. For example, the change in the number of weak relationships in the last five stages of 1046 is 0, 0, 0, 0, 1.

Table A5
ERT topics in stem cells research domain.

No.	***Topic id***	***Topic labels of stem cells***	***ERTs detail description***	***ERTs short description***
1	353	stem cell; cell; surface; mesenchymal stem; EXTRACELLULAR MATRIX; substrate; differentiation; tissue; hydrogel; biomaterial	Stem cells have a key role in the maintenance and regeneration of tissues and they are located in a specific microenvironment, defined as niche. Extracellular matrix (ECM) represents an essential player in stem cell niche, since it can directly or indirectly modulate the maintenance, proliferation, self-renewal and differentiation of stem cells. In the past, biochemical signals were the focus of induced stem cell differentiation. Biophysical signal mediated stem cell differentiation could also serve as an important inducer. Engineered biomaterials able to mimic the in vivo characteristics of stem cell niche provide suitable in vitro tools for dissecting the different roles exerted by the ECM and its molecular components on stem cell behavior.	Extracellular matrix could modulate stem cells' fate
2	2276	stem cell; intestinal stem; expression; organoid; cancer; intestinal; COLORECTAL CANCER; crypt; model; Lgr5	The intestinal epithelium is the most rapidly self-renewing tissue in adult mammals. It is currently believed that four to six crypt stem cells reside at the +4 position immediately above the Paneth cells in the small intestine. There is an exclusive expression of Lgr5 in cycling columnar cells at the crypt base. Epithelial organoids recapitulate multiple aspects of real organs, making them promising models of organ development, function and disease. The ability to differentiate hESCs or hiPSCs into colon stem cells could provide a platform to model patient-specific gastro-intestinal disease.	Stem cells in intestinal epithelium
3	1460	stem cell; scaffold; tissue engineering; cell; extracellular matrix; tissue; mesenchymal stem; VITRO; regeneration; decellularized	The potential of decellularized scaffolds, derived from native tissues or organs in the form of scaffolds has been evolved as a promising approach in tissue-regenerative medicine for translating functional organ replacements. In recent years, donor organs, such as heart, liver, lung and kidneys, have been reported to provide acellular extracellular matrix (ECM)-based scaffolds through the process called 'decellularization' and proved to show the potential of recellularization with selected cell populations, particularly with stem cells. In fact, decellularized stem cell matrix (DSCM) has also emerged as a potent biological scaffold for controlling stem cell fate and function during tissue organization.	Decellularized stem cell in tissue-regenerative medicine
4	142	DNA methylation; STEM CELL; epigenetic; gene; gene expression; EMBRYONIC STEM; expression; human; cell; development	DNA methylation is a stable epigenetic mark that can be inherited through multiple cell divisions. During development and cell differentiation, DNA methylation is dynamic, but some DNA methylation patterns may be retained as a form of epigenetic memory. DNA methylation profiles can be useful for the lineage classification and quality control of stem cells such as embryonic stem cells, induced pluripotent cells and mesenchymal stem cells.	DNA methylation in stem cells
5	60			

(*continued on next page*)

Table A5 (*continued*)

No.	*Topic id*	*Topic labels of stem cells*	*ERTs detail description*	*ERTs short description*
		stem cell; mesenchymal stem; ARTICULAR CARTILAGE; scaffold; tissue engineering; chondrocyte; chondrogenic differentiation; bone marrow; growth factor; cartilage repair	The articular cartilage is unique in that it contains only a single type of cell and shows poor ability for spontaneous healing. Currently, approaches for treating cartilage defects include surgical and nonsurgical approaches, as well as cartilage tissue engineering. For standard cartilage tissue engineering, three elements are required, i.e., a scaffold, growth factors, and seed cells. With advancements in stem cell research, the main sources of cells for cartilage tissue engineering are embryonic and mesenchymal stem cells, which have been shown to be promising alternatives in recent years. Under certain conditions, several types of stem cells, including embryonic stem cells, mesenchymal stem cells, induced pluripotent stem cells, and cartilage progenitor cells, showed potential for applications in chondrogenic differentiation.	Stem cells provide sources for cartilage tissue engineering
6	921	stem cell; dental pulp; expression; periodontal ligament; mesenchymal stem; pulp stem; osteogenic differentiation; human dental; VITRO; growth factor	The transplantation of replacement cells or tissues generated by the patient's own stem cells has been considered as a potential treatment for repairing bone defects. In particular, stem cells derived from the pulp of human exfoliated deciduous teeth and periodontal ligaments might be an ideal cell source for repairing bone defects in pediatric patients. Epidermal growth factor (EGF) and basic fibroblast growth factor (bFGF) play an important role in extracellular matrix mineralization, a complex process required for proper bone regeneration. Human dental pulp stem/progenitor cells (hDPSCs) in combination with EGF could be an effective stem cell-based therapy for bone tissue engineering applications in periodontics and oral implantology.	Human dental pulp stem/progenitor cells (hDPSCs) based therapy for bone tissue engineering
7	727	beta catenin; stem cell; Wnt signaling; expression; Wnt beta; signaling pathway; cell; catenin signaling; protein; gene	The canonical Wnt pathway, has a prominent role in stem cell proliferation and differentiation. Aberrant regulation has been shown to effect stem cell behavior and can lead to the development of cancer. During embryogenesis, stem cells demonstrate a requirement for β-catenin in mediating the response to Wnt signaling for their maintenance and transition from a pluripotent state. In adult stem cells, Wnt signaling functions at various hierarchical levels to contribute to specification of different tissues.	Regulation of Wnt signaling pathway in stem cells
8	1046	stem cell; EZH2; expression; cell; gene; protein; gene expression; differentiation; EMBRYONIC STEM; chromatin	Enhancer of zeste homolog 2 (EZH2), a catalytic component of polycomb repressive complex 2, serves as a histone methyltransferase toward histone H3K27 trimethylation and also recruits DNA methyltransferases to regulate gene expression and chromatin structure. EZH2 play critical roles in stem cell maintenance and cell fate decision in differentiation into specific cell lineages.	The role of histone methyltransferase EZH2 in stem cell fate
9	161	acute myeloid; myeloid leukemia; stem cell; myelodysplastic syndrome; cell transplantation; patient; AML; hematopoietic stem; leukemia AML; treatment	Haematopoietic stem cells (HSC) are situated at the apex of the haematopoietic differentiation hierarchy, ensuring the life-long supply of mature haematopoietic cells and forming a reservoir to replenish the haematopoietic system in case of emergency such as acute blood loss. To maintain a balanced production of all mature lineages and at the same time secure a stem cell reservoir, intricate regulatory programs have evolved to control multi-lineage differentiation and self-renewal in haematopoietic stem and progenitor cells (HSPCs). Leukaemogenic mutations commonly disrupt these regulatory programs causing a block in differentiation with simultaneous enhancement of proliferation. Increasing numbers of patients are receiving allogeneic hematopoietic cell transplantation (HCT) for acute myeloid leukaemia (AML).	Clinical application of hematopoietic stem cells
10	221	stem cell; retina; transplantation; cell; differentiation; retinal pigment; photoreceptor; pluripotent stem; MACULAR DEGENERATION; progenitor cell	Retinal degenerative diseases are the leading causes of blindness worldwide. Replacing lost retinal cells via stem cell-based therapies is an exciting, rapidly advancing area of translational research that has already entered the clinic. In age-related macular degeneration (AMD), stem cells could possibly replace or regenerate disrupted pathologic retinal pigment epithelium (RPE), and produce supportive growth factors and cytokines such as brain-derived neurotrophic factor.	Application of stem cells in retinal degenerative diseases
11	1290	acute lymphoblastic; lymphoblastic leukemia; patient; stem cell; cell transplantation; residual disease; minimal residual; treatment; cell precursor; relapse	Monitoring of minimal residual disease (MRD) has become routine clinical practice in frontline treatment of virtually all childhood acute lymphoblastic leukaemia (ALL) and in many adult ALL patients. MRD diagnostics has proven to be the strongest prognostic factor. Also in relapsed ALL patients and patients undergoing stem cell transplantation,	Clinical application of stem cells in acute lymphoblastic leukaemia

(*continued on next page*)

Table A5 (*continued*)

No.	*Topic id*	*Topic labels of stem cells*	*ERTs detail description*	*ERTs short description*
			MRD diagnostics is guiding treatment decisions. Multiple methods to detect MRD exist, each with their own advantages and disadvantages. Multiparameter flow cytometry and quantitative polymerase chain reaction are the most commonly used methods of MRD detection in clinical practice, although there is promise in the use of more sensitive assays utilizing next-generation sequencing that may be able to further refine MRD-based risk stratification.	
12	2	stem cell; microRNA; miRNA; expression; breast cancer; regulation; function; MESENCHYMAL TRANSITION; cancer cell; MicroRNAs miRNA	At present, research in the areas of stem-cell biology, RNA interference (RNAi), and microRNAs (miRNAs) is the source of immense scientific and medical interest. Despite their apparent differences, stem-cell biology, RNAi, and miRNA are increasingly being shown to have a great deal in common. Small RNAs, particularly those represented by miRNAs and RNAi, play a central role in the regulation of embryonic and adult stem-cell development. Recent research has revealed that in stem cells, miRNA- and RNAi-mediated gene regulation is one of the determinates controlling the state of cell differentiation, with the small RNAs serving as key elements involved in regulatory network control of pluripotent cell-fate determination. From a technological perspective, RNAi and miRNA-based methods are being developed both to study the processes of stem-cell differentiation and create clinically useful therapies for the treatment of disease and preservation of wellness.	Small RNAs play a central role in the regulation of stem cells
13	1199	stem cell; cell; development; CEREBRAL CORTEX; neuron; progenitor cell; expression; differentiation; neurogenesis; neural stem	Neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, and Huntington's disease currently affect tens of millions of people worldwide. A growing number of pre-clinical studies have suggested that transplantation of neural stem cells (NSCs) could offer a promising new therapeutic approach for neurodegeneration.	Application of neural stem cells in the treatment of neurodegenerative disorders
14	710	stem cell; cancer stem; breast cancer; cancer cell; expression; CSC; INITIATING CELL; tumor; CD133; self-renewal	Cancer stem cells (CSCs) have emerged as an underlying cause of cancer relapse and resistance to treatment. Several CSC markers have been identified from many types of tumours, and these markers are also being used for isolation and enrichment of CSCs. Cluster of differentiation CD133 is a well-characterized CSC marker, and it is involved in tumor cell proliferation, metastasis, tumorigenesis, and recurrence, as well as chemo- and radio-resistance.	The biology of cancer stem cells
15	814	spinal cord; cord injury; stem cell; FUNCTIONAL RECOVERY; transplantation; SCI; mesenchymal stem; neural stem; injury SCI; bone marrow	Spinal cord injury (SCI) is a catastrophic nervous system disorder. Following a SCI, motor dysfunction will persist unless the injured region is recovered. However, neurogenesis is an un- common phenomenon in the central nervous system (CNS) and self-healing potentials in injured cells are extremely limited. There could be an alternative way to replenish damaged cells following SCI: stem cell therapy. Stem cells have a considerable ability to renew themselves and differentiate into any desired type of cell. Their capability to generate all kinds of cells provides an opportunity to replenish damaged and dysfunctional cells. Cell therapy can regenerate neural connections which in turn leads to improved functional recovery.	Stem cell therapy for spinal cord injury
16	254	stem cell; mesenchymal stem; stromal cell; bone marrow; MSC; mesenchymal stromal; cells MSC; adipose tissue; expression; derived mesenchymal	Human MSCs, that represent a mesoderm derived population of progenitors, are easily expanded in culture. They are capable to differentiate into osteoblasts, chondrocytes, and adipocytes and exhibit the potential to repair or regenerate damaged tissues. The best characterized source of human MSCs to date is the bone marrow; recently, foetal sources, such as amniotic fluid, umbilical cord, amniotic membranes, or placenta, have also attracted increased attention. Thus, MSCs may represent a valuable tool for tissue repair and cell therapeutic applications.	Research on mesenchymal stem cells and its application
17	581	stem cell; mesenchymal stem; scaffold; bone regeneration; growth factor; morphogenetic protein; tissue engineering; osteogenic differentiation; bone formation; BMP	Bone tissue engineering uses the principles and methods of engineering and life sciences to study bone structure, function and growth mechanism for the purposes of repairing, maintaining and improving damaged bone tissue. To regenerate bone structures which can be used for treatment of patients, bone growth must be induced by a number of bioactive implantable materials, cell types and intracellular, and extracellular molecular signaling pathways. Since mesenchymal stem cells and their	Regeneration and repair of bone

(*continued on next page*)

Table A5 (*continued*)

No.	Topic id	Topic labels of stem cells	ERTs detail description	ERTs short description
			differentiation during remodeling processes have important roles in bone regeneration, it is believed that understanding molecular signaling pathways involved is crucial to the development of bone implants, bone substitute materials, and cell-based scaffolds for bone regeneration.	
18	1142	stem cell; hair follicle; skin; cell; expression; differentiation; regeneration; hair; follicle stem; keratinocyte	Various types of stem cells reside in the skin, including keratinocyte progenitor cells, melanocyte progenitor cells, skin-derived precursors (SKPs), and nestin-expressing hair follicle-associated-pluripotent (HAP) stem cells. HAP stem cells, located in the bulge area of the hair follicle, have been shown to differentiate to nerve cells, glial cells, keratinocytes, smooth muscle cells, cardiac muscle cells, and melanocytes. HAP stem cells are positive for the stem-cell marker CD34, as well as K15-negative, suggesting their relatively undifferentiated state. Therefore, HAP stem cells may be the most primitive stem cells in the skin. Moreover, HAP stem cells can regenerate the epidermis and at least parts of the hair follicle. These results suggest that HAP stem cells may be the origin of other stem cells in the skin.	Skin-resident stem cells
19	461	multiple myeloma; stem cell; CELL TRANSPLANTATION; patient; treatment; myeloma MM; bone marrow; therapy; autologous stem; patients multiple	Autologous stem cell transplantation (ASCT) has long been considered frontline therapy for newly diagnosed myeloma patients. A combination of novel drugs with ASCT in a sequential treatment approach can attain long-term survival and perhaps cure a subset of patients. ASCT will likely remain an important platform to develop curative strategies in the foreseeable future.	Clinical application of autologous stem cell transplantation
20	261	patient; stem cell; Hodgkin lymphoma; cell transplantation; cell lymphoma; NON HODGKIN; autologous stem; treatment; high dose; relapsed refractory	Despite the high first-line cure rates in patients with Hodgkin Lymphoma (HL) still 10%−20% of patients suffer from relapsed or refractory disease. The introduction of high-dose (HD) chemotherapy (CT) and autologous stem cell (ASCT) or bone marrow transplant (ABMT) in the last two decades has improved the prognosis of patients with refractory or relapsed Hodgkin lymphoma (HL) over conventional-dose salvage CT.	Stem cell therapy for Hodgkin's lymphoma
21	867	stem cell; Drosophila; differentiation; protein; development; SELF RENEWAL; gene; expression; niche; tissue	The discovery of Drosophila stem cells with striking similarities to mammalian stem cells has brought new hope for stem cell research. Genes that controlled Drosophila stem cells were discovered to have functional homologs in mammalian stem cells. Compared to mammalian model organisms, it is easy and inexpensive to work with Drosophila. Furthermore, shorter generation time, small size, and fewer ethical issues makes Drosophila an attractive animal model. Like mammals, Drosophila germline stem cells (GSCs) are controlled by both intrinsic as well as external signals. Several factors including BMP signals were found to be indispensable for sustaining GSCs in Drosophila. Asymmetric division of GSCs to produce and maintain a daughter GSC is regulated by gene expression in adjacent somatic cells also.	Drosophila as a model to study adult stem cells
22	648	stem cell; limbal stem; ocular surface; epithelial cell; amniotic membrane; cornea; corneal; expression; transplantation; limbal epithelial	Stem cells are self-renewing cells that can differentiate into specialized cell type(s). Pluripotent stem cells, i.e. embryonic stem cells (ESC) or induced pluripotent stem cells (iPSC) differentiate into cells of all three embryonic lineages. Multipotent stem cells, like hematopoietic stem cells (HSC), can develop into multiple specialized cells in a specific tissue. Unipotent cells differentiate only into one cell type, like e.g. satellite cells of skeletal muscle. There are many examples of successful clinical applications of stem cells. Skin stem cells are used to heal severe burns, while limbal stem cells (LSC) can regenerate the damaged cornea. The cornea epithelium represents one of the most rapidly regenerating mammalian tissues, undergoing full-turnover over the course of approximately 1–2 weeks.	Concept and application of limbal stem cells
23	1095	skeletal muscle; stem cell; satellite cell; differentiation; muscle regeneration; regeneration; cell; muscle stem; expression; myogenesis	Adult skeletal muscle in mammals is a stable tissue under normal circumstances but has remarkable ability to repair after injury. Skeletal muscle regeneration is a highly orchestrated process involving the activation of various cellular and molecular responses. As skeletal muscle stem cells, satellite cells play an indispensable role in this process. The self-renewing proliferation of satellite cells not only maintains the stem cell population but also provides numerous myogenic cells, which proliferate, differentiate, fuse, and lead to new myofiber formation and reconstitution of a functional contractile apparatus.	Adult stem cells and skeletal muscle regeneration

(*continued on next page*)

Table A5 (*continued*)

No.	*Topic id*	*Topic labels of stem cells*	*ERTs detail description*	*ERTs short description*
24	107	cell; stem cell; patient; immunotherapy; cancer; antigen receptor; chimeric antigen; tumor; therapy; expression	Human pluripotent stem cells (PSCs) provide a promising resource to produce immune cells for adoptive cellular immunotherapy to better treat and potentially cure otherwise lethal cancers. Cytotoxic T cells and natural killer (NK) cells can now be routinely produced from human PSCs. These PSC-derived lymphocytes have phenotype and function similar to primary lymphocytes isolated from peripheral blood. PSC-derived T and NK cells have advantages compared with primary immune cells, as they can be precisely engineered to introduce improved anti-tumor activity and produced in essentially unlimited numbers. Chimeric antigen receptors (CARs) are engineered molecules that can be introduced into T cells to enable them to target specific tumor antigens. CAR T cells targeting CD19 have shown promise in patients with relapsed and refractory B-cell neoplasms, including those with acute lymphoblastic leukaemia, chronic lymphocytic leukaemia, and non-Hodgkin lymphomas.	Application of stem cells in cellular immunotherapy
25	965	stem cell; germ cell; spermatogonial stem; differentiation; expression; spermatogenesis; SSC; vitro; mice; Sertoli cell	Spermatogenesis is a continuous and productive process supported by the self-renewal and differentiation of spermatogonial stem cells (SSCs), which arise from undifferentiated precursors known as gonocytes and are strictly controlled in a special 'niche' microenvironment in the seminiferous tubules. Maintenance of SSCs and induction of spermiogenesis in vitro may provide a therapeutic strategy to treat male infertility. Sertoli cells, the only somatic cell type in the tubules, directly interact with SSCs to control their proliferation and differentiation through the secretion of specific factors.	Regulation of spermatogonial stem cells and spermatogenesis
26	469	stem cell; liver fibrosis; expression; progenitor cell; stellate cell; mesenchymal stem; hepatic stellate; hepatocyte; liver disease; bone marrow	The high regenerative potential of the liver is driven by parenchymal and non-parenchymal cells, which restore the original liver mass after injury by cell proliferation. The contribution of stem- and progenitor cells to liver regeneration is mainly observed when hepatocyte proliferation is impaired. However, the origin of stem/progenitor cells and their effectivity to restore injured liver is currently discussed controversially. Hepatic stellate cells, which are mainly known for their contribution to fibrosis in chronic liver diseases, were recently identified as mesenchymal stem cells of the liver. Stellate cells are also involved in liver regeneration and fulfill a dual function by supporting neighbouring cells and developing into liver epithelial cells.	Stem cells in liver regeneration

References

Apreda, R., Bonaccorsi, A., dell'Orletta, F., Fantoni, G., 2019. Expert forecastand realized outcomes in technology foresight. Technol. Forecast. Soc. Change 141, 277–288.

Arora, S.K., Youtie, J., Shapira, P., Gao, L., Ma, T., 2013. Entry strategies in an emerging technology: a pilot web-based study of graphene firms. Scientometrics 95, 1189–1207. https://doi.org/10.1007/s11192-013-0950-7.

Bai, G., Zheng, Y., Wu, X., Jin, J., Liu, Q., 2017. Research and demonstration on forecasting method of disruptive technology based on literature knowledge correlation. J. Intell. 36 (9), 48-44.

Bioon, 2018. Research and application of stem cells and new materials in regenerative medicine. 2018 (10th) International Conference on Stem Cells. https://meeting.bioon.com/2018StemCells/news-detail/b04d4462ec057124. China 2018-11-09-2018-11-10•(2019-9-16).

Bolger, F., Wright, G., 2017. Use of expert knowledge to anticipate the future: issues, analysis and directions. Int. J. Forecast. 33, 230–243.

Bostock, M., Ogievetsky, V., Heer, J., 2011. D^3 data-driven documents. IEEE Trans. Visualizat. Comput. Graph. 12, 2301–2309.

Boyack, K., Glänzel, W., Gläser, J., et al. (February 2017). Topic extraction challenge. http://www.topic-challenge.info/.

Boyack, K.W., Klavans, R., Small, H., Ungar, L., 2014. Characterizing the emergence of two nanotechnology topics using a contemporaneous global micro-model of science. J. Eng. Technol. Manag. 32, 147–159.

Burmaoglu, S., Sartenaer, O., Porter, A., Li, M., 2019. Analysing the theoretical roots of technology emergence: an evolutionary perspective. Scientometrics 119 (1), 97–118.

Carley, S.F., Newman, N.C., Porter, A.L., Garner, J.G., 2018. An indicator of technical emergence. Scientometrics 115, 35–49. https://doi.org/10.1007/s11192-018-2654-5.

Chen, C. (2001). A new introduction to natural dialectics Shenyang: northeastern University press.

Chen, C., 2006. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. J. Am. Soc. Inf. Sci. Technol. 57 (3), 359–377.

Chen, C., 2012. Turning Points: The Nature of Creativity. Springer Science & Business Media.

14 ClinicalTrials.gov.https://clinicaltrials.gov/ct2/results?cond=stem+cell+cord+injury&term=&cntry=&state=&city=&dist=. [EB/OL]. [2019-9-16].

Cozzens, S., Gatchair, S., Kang, J., Kim, K., Lee, H.J., Ordóñez, G., Porter, A., 2010. Emerging technologies: quantitative identification and measurement. Technol. Anal. Strategic Manag. 22 (3), 361–376.

CURTIS, E., MARTIN J, R., GABEL, B., et al., 2018. A first-in-human, phase I study of neural stem cell transplantation for chronic spinal cord Injury. Cell Stem. Cell., 2018 22 (6), 941–950.

Dahlin, K.B., Behrens, D.M., 2005. When is an invention really radical?: defining and measuring technological radicalness. Res. Policy 34 (5), 717–737.

18 Day, G.S., & Schoemaker, P.J.H. (2006). Peripheral Vision: Detecting the Weak Signals That Will Make or Break Your Company (Boston MA: Harvard Business School Press). BUSINESS WARGAMING, 128.

De Haan, J., 2006. How emergence arises. Ecol. Complex. 3 (4), 293–301.

Eck, N.J.v., & Waltman, L. (2019). VOSviewer Manual. Retrieved from http://www.vosviewer.com/.

Gao, J.P., Ding, K., Teng, L., Pang, J., 2012. Hybrid documents co-citation analysis: making sense of the interaction between science and technology in technology diffusion. Scientometrics 93 (2), 459–471.

Gerrish, S.M., Blei, D.M., 2010. A language-based approach to measuring scholarly impact. Proceedings of the 27th International Conference on Machine Learning. Omnipress, pp. 375–382.

Glänzel, W., Thijs, B., 2012. Using 'core documents' for detecting and labelling new emerging topics. Scientometrics 91 (2), 399–416.

Goldstein, J., 1999. Emergence as a construct: history and issues. Emergence 1 (1), 49–72.

Granovetter, M., 1983. The strength of weak ties: a network theory revisited. Sociol. theory 201–233.
Guo, H., Weingart, S., Börner, K., 2011. Mixed-indicators model for identifying emerging research areas. Scientometrics 89 (1), 421–435.
Guyette J, P., Charest J, M., Mills R, W., et al., 2016. Bioengineering human myocardium on native extracellular matrix. Circulat. Res., 2016 118 (1), 56–72.
Hansmann, T., Niemeyer, P., 2014. Big data-characterizing an emerging research field using topic models. In: 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 1. IEEE, pp. 43–51.
Hido, S., Suzuki, S., Nishiyama, R., Imamichi, T., Takahashi, R., Nasukawa, T., Ueno, T., 2012. Modeling patent quality: a system for large-scale patentability analysis using text mining. Inf. Media Technol. 7 (3), 1180–1191.
Hiromi, K., Ken, K., Adler, A.F., 2018. Generation and post-injury integration of human spinal cord neural stem cells. Nature Methods, 2018 15 (9), 723–731.
Hu Z.Y., Zeng R.Q., Qin X.C., Wei L., Zhang Z. (2018) A method of biomedical knowledge discovery by literature mining based on SPO predications: a case study of induced pluripotent stem cells. In: Perner P. (eds) Machine Learning and Data Mining in Pattern Recognition. MLDM 2018. Lecture Notes in Computer Science, vol 10935. Springer, Cham.
Hu, C.P., Hu, J.M., Deng, S.L., Liu, Y., 2013. A co-word analysis of library and information science in China. Scientometrics 97 (2), 369–382.
Huang, M.H., Chang, C.P., 2014. Detecting research fronts in OLED field using bibliographic coupling with sliding window. Scientometrics 98 (3), 1721–1744.
Huang, M.H., Yang, H.W., Chen, D.Z., 2015. Increasing science and technology linkage in fuel cells: a cross citation analysis of papers and patents. J Informetr 9 (2), 237–249.
Huang, Y., Porter, A.L., Zhang, Y., Lian, X., Guo, Y., 2019. An assessment of technology forecasting: revisiting earlier analyses on dye-sensitized solar cells (DSSCs). Technol Forecast Soc Change 146, 831–843.
Huang, Y., Zhu, D., Qian, Y., Zhang, Y., Porter, A.L., Liu, Y., Guo, Y., 2017. A hybrid method to trace technology evolution pathways: a case study of 3D printing. Scientometrics 111 (1), 185–204.
Hung, S.-.C., Chu, Y.-.Y., 2006. Stimulating new industries from emerging technologies: challenges for the public sector. Technovation 26 (1), 104–110.
Iwami, S., Mori, J., Sakata, I., Kajikawa, Y., 2014. Detection method of emerging leading papers using time transition. Scientometrics 101 (2), 1515–1533.
Iyiou. (2019). Research progress and industry analysis of stem cell therapy. https://www.iyiou.com/p/95934.html. [2019-9-16].
Jarić, I., Knežević-Jarić, J., Lenhardt, M., 2014. Relative age of references as a tool to identify emerging research fields with an application to the field of ecology and environmental sciences. Scientometrics 100 (2), 519–529.
Joung, J., Kim, K., 2017. Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data. Technol. Forecast. Soc. Change 114, 281–292.
Klavans, R., Boyack, K.W., 2017. Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? J. Assoc. Inf. Sci. Technol. 68 (4), 984–998.
Kwon, H., Kim, J., Park, Y., 2017. Applying LSA text mining technique in envisioning social impacts of emerging technologies: the case of drone technology. Technovation 60, 15–28.
Lee, C., Kwon, O., Kim, M., Kwon, D., 2018. Early identification of emerging technologies: a machine learning approach using multiple patent indicators. Technol. Forecast. Soc. Change 127, 291–303.
Lee, W., 2008. How to identify emerging research fields using scientometrics: an example in the field of Information Security. Scientometrics 76 (3), 503–525.
Leydesdorff, L., 1987. Various methods for the mapping of science. Scientometrics 11 (5–6), 295–324.
Li, M., 2018. Visualizing Dispersed Risk Signals for a Specific emerging technology: a novel approach of keywords aggregation across topics (KAaT. Paper presented at the Global TechMining Conference 2018.
Li, M., Porter, A.L., Suominen, A., 2018. Insights into relationships between disruptive technology/innovation and emerging technology: a bibliometric perspective. Technol. Forecast. Soc. Change 129, 285–296.
The delphi method. In: Linstone, H.A., Turoff, M. (Eds.), 1975. Reading. Wesley, MA: Addison, pp. 3–12.
Liu, Y., 2007. Quality estimation of patent based on text mining and its empirical research. Comput. Eng. Appl. 43, 12–14.
Liu, Z., Yin, Y., Liu, W., Dunford, M., 2015. Visualizing the intellectual structure and evolution of innovation systems research: a bibliometric analysis. Scientometrics 103 (1), 135–158.
López-Menéndez, A.J., Pérez-Suárez, R., 2019. Acknowledging uncertainty in economic forecasting. Some insight from confidence and industrial trend surveys. Entropy 21 (4), 413.
López-Menéndez, A.J., & Pérez-Suárez, R. (2020). Entropy application for forecasting.
Luo, R., Xu, H., Dong, K., 2018. A Study on the main recognition methods of frontier research and its comparative study. Lib. Inf. Serv. 62 (23), 119–131.
Mairesse, J., Pezzoni, M., 2018. Novelty in Science: the impact of French physicists' novel articles. 23rd International Conference on Science and Technology Indicators (STI 2018), September 12-14, 2018. Centre for Science and Technology Studies (CWTS), Leiden, The Netherlands.
Martin, B.R., 1995. Foresight in science and technology. Technol. Anal. Strategic Manag. 7 (2), 139–168.
Meyer, M., Debackere, K., Glänzel, W., 2010. Can applied science be 'good science'? Exploring the relationship between patent citations and citation impact in nanoscience. Scientometrics 85 (2), 527–539.
Min C., Bu Y., & Sun J. (2019). Comparing breakthrough and non-breakthrough papers from early citing structures.
Mitchell, S.D., 2007. The import of uncertainty. The Pluralist 2 (1), 58–71.
Narin, F., Noma, E., 1985. Is technology becoming science. Scientometrics 7 (3–6), 369–381.
Porter, A.L., Detampel, M.J., 1995. Technology opportunities analysis. Technol. Forecast. Soc. Change 49 (3), 237–255.
Porter, A.L., Garner, J., Carley, S.F., Newman, N.C., 2019. Emergence scoring to identify frontier R&D topics and key players. Technol. Forecast. Soc. Change 146, 628–643.
Porter, A.L., Roessner, J.D., Jin, X.Y., Newman, N.C., 2002. Measuring national 'emerging technology'capabilities. Sci. Pub. Policy 29 (3), 189–200.
Qi, Y., 2015. Study on the patent novelty and its measurement. National Science Library. Chinese Academy of Sciences.
Reardon, S., 2014. Text-mining offers clues to success: US intelligence programme analyses language in patents and papers to identify next big technologies. Nature 509 (7501), 410–411.
Rhee, M., 2004. Network updating and exploratory learning environment. J. Manag. Stud. 41 (6), 933–949.
Roche, I., Besagni, D., François, C., Hörlesberger, M., Schiebel, E., 2010. Identification and characterisation of technological topics in the field of Molecular Biology. Scientometrics 82 (3), 663–676.
Roper, A.T., Cunningham, S.W., Porter, A.L., Mason, T.W., Rossini, F.A., Banks, J., 2011. Forecasting and Management of Technology, 2d Edition. John Wiley, New York.
Rotolo, D., Hicks, D., Martin, B.R., 2015. What is an emerging technology? Res. Policy 44 (10), 1827–1843.
Santa Soriano, A, Lorenzo Álvarez, C, Torres Valdés, R.M, 2018. Bibliometric analysis to identify an emerging research area: public Relations Intelligence—A challenge to strengthen technological observatories in the network society. Scientometrics 115, 1591–1614.
Schmoch, U., 1993. Tracing the knowledge transfer from science to technology as reflected in patent indicators. Scientometrics 26 (1), 193–211.
Schoemaker, P.J., Day, G.S., Snyder, S.A., 2013. Integrating organizational networks, weak signals, strategic radars and scenario planning. Technol. Forecast. Soc. Change 80 (4), 815–824.
Schumpeter, J.A., 1939. Business Cycles, 1. McGraw-Hill, New York.
Sheen, J., & Wang, B.Z. (2017). Estimating macroeconomic uncertainty from surveys-A mixed frequency approach. Available at SSRN 3020697.
Stahl, B.C., 2011. What does the future hold? A critical view of emerging information and communication technologies and their social consequences. Researching the Future in Information Systems. Springer, Berlin, Heidelberg, pp. 59–76.
Small, H., 1973. Co-citation in the scientific literature: a new measure of the relationship between two documents. J. Am. Soc. Inf. Sci. 24 (4), 265–269.
Small, H., Griffith, B., 1974. The structure of scientific literatures I: identifying and graphing specialties. Sci. Stud. (St. Bonaventure) 4 (1), 17–40.
Small, H., Boyack, K.W., Klavans, R., 2014. Identifying emerging topics in science and technology. Res. Policy 43 (8), 1450–1467.
Srinivasan, R., 2008. Sources, characteristics and effects of emerging technologies: research opportunities in innovation. Ind. Market. Manag. 37 (6), 633–640.
Stein, B., Eissen, S.M.Z., 2004. Topic identification: framework and application. IEEE Internet Comput. 11 (3), 67–73.
Stirling, A., 2007. Risk, precaution and science: towards a more constructive policy debate: talking point on the precautionary principle. EMBO Rep. 8 (4), 309–315.
Sugimoto, C.R., Larivière, V., 2018. Measuring Research: What Everyone Needs to Know. Oxford University Press.
Sullivan W, J., Mullen P, J., Schmid E, W., et al., 2018. Extracellular matrix remodeling regulates glucose metabolism through TXNIP destabilization. Cell 175 (1), 117–132, 2018.
Takeda, Y., Kajikawa, Y., 2009. Optics: a bibliometric approach to detect emerging research domains and intellectual bases. Scientometrics 78 (3), 543–558.
Traag, V., Waltman, L., & van Eck, N.J. (2018). From Louvain to Leiden: guaranteeing well-connected communities. arXiv preprint arXiv:1810.08473.
Tu, Y.-.N., Seng, J.-.L., 2012. Indices of novelty for emerging topic detection. Inf. Process. Manag. 48, 303–325.
Velden, T., Boyack, K.W., Gläser, J., Koopman, R., Scharnhorst, A., Wang, S., 2017. Comparison of topic extraction approaches and their results. Scientometrics 111 (2), 1169–1221.
Villaseñor Terán, M, 2017. Philosophical Explorations for a Concept of Emerging Technologies. Research in Philosophy and Technology, Techné.
Waltman, L., Van Eck, N.J., 2012. A new methodology for constructing a publication-level classification system of science. J. Am. Soc. Inf. Sci. Technol. 63 (12), 2378–2392.
Waltman, L., Van Eck, N.J., 2013. A smart local moving algorithm for large-scale modularity-based community detection. Eur. Phys. J. B 86 (11), 471.
Wang (2016). Recurrent/refractory Hodgkin lymphoma autologous hematopoietic stem cell transplantation before treatment with bendamustine, gemcitabine, and vinorelbine as rescue chemotherapy. https://www.bjcancer.org/Mobile/Article/Index/4091. (2019-9-16).
Wang, Q., 2018. A bibliometric model for identifying emerging research topics. J. Assoc. Inf. Sci. Technol. 69 (2), 290–304.
Wei, L., Xu, H., Wang, Z., Dong, K., Wang, C., Fang, S., 2017. Topic detection based on weak tie analysis: a case study of LIS research. J. Data Inf. Sci. 1 (4), 81–101.
Weismayer, C., Pezenka, I., 2017. Identifying emerging research fields: a longitudinal latent semantic keyword analysis. Scientometrics 113 (3), 1757–1785.
Winnink, J., 2018. Science-technology interactions: using NPLRs as glue. Paper presented at the Global TechMining Conference 2018.
Xu, H., Winnink, J., Yue, Z., Liu, Z., Yuan, G., 2020a. Topic-linked innovation paths in science and technology. J. Informetr. 14 (2), 101014.

Xu, H., Yue, Z., Wang, C., Dong, K., Pang, H., Han, Z., 2017. Multi-source data fusion study in scientometrics. Scientometrics 111 (2), 773–792.
Xu, S, Liu, J., Zhai, D., An, X., Wang, Z., Pang, H., 2018. Overlapping Thematic structures extraction with mixed-membership stochastic blockmodel. Scientometrics 117 (1), 61–84.
Xu, S., Hao, L., An, X., Pang, H., Li, T., 2019b. Review on emerging research topics with key-route main path analysis. Scientometrics 1–18.
Xu, S., Hao, L., An, X., Yang, G., Wang, F., 2019c. Emerging research topics detection with multiple machine learning models. J. Informetr. 13 (4), 100983.
Yan, E., 2014. Research dynamics: measuring the continuity and popularity of research topics. J. Informetr. 8 (1), 98–110.
Yaqub, O., 2018. Serendipity: towards a taxonomy and a theory. Res. Policy 47 (1), 169–179.
Youzhunet. (2019) Trend 37 stem cell clinical projects have passed the national record and have a very wide range of disease treatment with full list. http://www.sohu.com/a/317270049_827862. [2019-9-16].
Zhang Y. (2011). Research on the scientific front detection by low-frequency occurrence phenomenon: Beijing: Chinese Academy of Sciences.
Zhang, Y., Huang, Y., Porter, A.L., Zhang, G., Lu, J., 2019. Discovering and forecasting interactions in big data research: a learning-enhanced bibliometric study. Technol Forecast Soc Change 146, 795–807.
Zitt, M., Lelu, A., Bassecoulard, E., 2011. Hybrid citation-word representations in science mapping: portolan charts of research fields? J. Am. Soc. Inf. Sci. Technol. 62 (1), 19–39.