Pharmaceutical R&D Portfolio Optimization with Minimum Borrowed Capital Based on Fuzzy Set Theory

Ming-Yu Zhang (zhang.mingyu2010@hotmail.com)
Third Affiliated Hospital of Southern Medical University

Yong-Jun Liu
South China Unviersity of Technology

Research Article

Keywords: Pharmaceutical project, R&D portfolio selection, minimum borrowed capital, fuzzy programming model, artificial bee colony algorithm

Posted Date: July 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-619187/v1

Noname manuscript No.
(will be inserted by the editor)

Pharmaceutical R&D portfolio optimization with minimum borrowed capital based on fuzzy set theory

Mingyu Zhang · Yong-Jun Liu*

Received: date / Accepted: date

Abstract Due to long lead times, uncertain outcomes and lack of enough historical data, pharmaceutical research and development (R&D) portfolio selection is a often very complex decision issue. The aim of this paper is to investigate pharmaceutical R&D portfolio selection with unavailable and unreliable project information, where the borrowed capital is allowed. Based on fuzzy set theory, we propose two pharmaceutical R&D portfolio optimization models with minimum borrowed capital by taking into account corporate strategy in developing new products, scarcity of resources, lack of investment budget and cardinality constraint. In the two proposed models, the pharmaceutical R&D company is assumed to achieve the objectives of maximizing terminal wealth and minimizing the cumulative borrowed capital over the whole investment horizon. Then, we transform the two proposed bi-objective models into the corresponding single-objective models by using the weighted sum approach and employ the modified artificial bee colony (ABC) algorithm to solve the transformed models. Finally, we provide a numerical example to illustrate the application of the proposed models.

Keywords Pharmaceutical project · R&D portfolio selection · minimum borrowed capital · fuzzy programming model · artificial bee colony algorithm

1 Introduction

R&D project portfolio selection is an organizational decision making task commonly found in organizations like government-funding agencies, universities, research in-

Mingyu Zhang
Department of Cardiovascular Medicine, The Third Affiliated Hospital of Southern Medical University, Guangzhou, Guangdong, 510063, PR China
E-mail: zhangmy_nysy@smu.edu.cn, zhang.mingyu2010@hotmail.com (M.Y.Zhang)

Yong-Jun Liu is the corresponding author
School of Business Administration, South China University of Technology, Guangzhou, 510641, P.R. China
E-mail: yjlgx0202@126.com; bmyjliu@scut.edu.cn (Y.J.Liu)

stitutes and technology-intensive firms (Tian et al., 2005). Along with the economic development and globalized marketplace, social competition becomes more and more fierce. As a result, the investment markets are changing faster and faster. Especially, global pharmaceutical R&D has witnessed an upsurge development, which is strongly motivated by highly unsatisfied medical demands, such as in cardiovascular, neoplastic and autoimmune diseases. Once approved on the market, a new drug's profit is prosperous. However, almost all the potential drug candidates face unpredictable performance with technical and market risks. Meanwhile, pharmaceutical R&D is a tough process characterized by long lead times and huge amount of investment. Over the past few decades, the expenditure on R&D has been increasing dramatically year by year. According to Mikulic (2020), the global pharmaceutical R&D expenditure in 2019 closed to 186 billion U.S. dollars. It is estimated that worldwide pharmaceutical R&D expenditure in 2022 will exceed 200 billion U.S. dollars. Actually, R&D has become a survival tool for any pharmaceutical company striving to achieve and maintain a strong competitive position in the future due to patent expiration of previous innovator drugs. The aim of R&D project portfolio selection is to select a subset of promising projects to construct a portfolio from a set of candidate projects based on multiple decision criteria. R&D project portfolio selection is always constrained by limited resources such as budget, research staff, laboratory space, and other technical scarcities. In many cases, organizations are often forced to pick out a certain number of projects from all candidate projects by means of project portfolio techniques mitigating the corresponding risks and enhancing the overall value of portfolio. Therefore, R&D project investment is a risky venture.

There are numerous decision models and methods that address R&D portfolio selection under the framework of probability theory. Among them, the investors are assumed to be able to produce precise values for the future attributes of R&D projects. For example, Abbassi et al. (2014) proposed a 0-1 nonlinear integer programming model for balancing portfolio values and risks of R&D project portfolio. Nishimura and Okada (2014) examined how R&D portfolios of drug pipelines affected pharmaceutical licensing, controlling firm size, diversity, and competition. Arratia et al. (2016) studied a static R&D project portfolio selection in public organizations. Beaujon et al. (2001) presented an R&D project portfolio selection model with a wide variety of constraints (e.g., capital, headcount, strategic intent, etc.). Bistline (2016) analyzed energy R&D decisions with uncertainty in research outcomes and markets. Çağlar and Gürel (2017) studied a public R&D project portfolio selection with project cancellations. Chen and Zhu (2011) proposed an integrated approach including nonparametric efficiency analysis, bootstrapping, and mean-variance optimization model to resource allocation problems, and applied it to R&D project budgeting. Gemici-Ozkan and Wu (2010) introduced a decision-support framework for the R&D portfolio selection problem faced by a major U.S. semiconductor manufacturer. Girotra et al. (2007) conducted an event study around the failure of phase III clinical trials and their effect on the market valuation of the firm. Hassanzadeh et al. (2014a) developed a robust optimization model to assist contract research organizations in making their primary business decision, i.e., selection and scheduling of new drug development project opportunities. Hassanzadeh et al. (2014b) presented a multi-objective binary integer programming model for R&D project portfolio selection with

competing objectives and uncertain coefficients in both objective functions and constraints. Huchzermeier and Loch (2001) developed a simple real option model for R&D project selection. Ringuest et al. (2000) described a methodology for the selection of R&D projects to add to or remove from an existing R&D portfolio. Rogers et al. (2002) proposed a stochastic optimization model for pharmaceutical R&D project portfolio selection by using a real options approachs. Solak et al. (2010) presented a multistage stochastic programming model for R&D project portfolios under endogenous uncertainty. Stummer and Heidenberger (2003) described a three-phase approach to assist R&D managers in obtaining the most attractive project portfolio. Tohumcu and Karasakal (2010) developed an approach based on analytic network process (ANP) and data envelopment analysis (DEA) to evaluate the performance of R&D projects. Pennings and Sereno (2011) presented a compound option valuation approach for pharmaceutical R&D project portfolio with technical and economic uncertainties. Karasakal and Aker (2017) developed multiple criteria sorting methods based on data envelopment analysis to evaluate R&D projects. Wang et al. (2018) established a two-stage stochastic programming model for portfolio optimization of R&D on low-carbon energy technology.

Notice that R&D project investment decision is often full of changes due to long lead times, uncertain outcomes and lack of enough historical data. In many cases, the collected data information for R&D project management are often at best uncertain and at worst very unreliable. Fuzzy set theory in Zadeh (1965) has been proved to be a powerful tool to deal with imprecise and uncertain data in many applications, which can also been used to characterize the uncertain information on project parameters. Some researchers have investigated R&D project selection problems by using fuzzy set theory to capture and model the uncertain project information. For example, Coffin and Taylor (1996) proposed a multiple criteria model for R&D project selection and scheduling using fuzzy logic. Mohanty et al. (2005) formulated a fuzzy ANP (analytic network process) model for R&D project selection. Wang et al. (2005) proposed a fuzzy AHP model for evaluating R&D projects from multiple disciplines. Carlsson et al. (2007) developed a fuzzy mixed integer programming model for R&D portfolio selection, where future cash flows were characterized by trapezoidal fuzzy number. Wang and Hwang (2007) used fuzzy set theory to model uncertain and flexible project information, and presented a fuzzy zero-one integer programming model for R&D project portfolio selection. Huang et al. (2008) proposed a fuzzy AHP method for government-sponsored R&D project selection. Sun et al. (2008) presented a group decision support approach to evaluate experts for R&D project selection. Chen and Hung (2010) developed an integrated fuzzy evaluation method for selecting a suitable outsourcing manufacturing partners in pharmaceutical R&D. Bhattacharyya et al. (2011) presented a fuzzy tri-objective programming model for R&D project portfolio selection. Collan and Luukka (2014) presented four new fuzzy similarity measure-based TOPSIS variants for R&D project ranking and evaluation. Hassanzadeh et al. (2012) adopted fuzzy pay-off method for R&D project evaluation and developed a fuzzy R&D project portfolio model. Biancardi and Villani (2017) presented a fuzzy approach to value R&D investments. Hesarsorkh et al. (2021) formulated a robust possibilistic optimization approach for pharmaceutical R&D project

portfolio selection and scheduling under uncertainty. It is important to remark that all the studies above do not account for R&D project portfolio with borrowed capital.

In the real world, due to the incomplete and imprecise information, investors may be forced to borrow a certain capital to make up budget shortfalls for the selected R&D projects. So, it is necessary to investigate fuzzy R&D project portfolio selection problem, where the borrowed capital is allowed in any intermediate or the final time period. To the best of our knowledge, no existing literature has addressed R&D project portfolio selection with borrowed capital in fuzzy environment. An integration of the borrowed capital over bankruptcy in R&D project portfolio is evidently needed. For this, we propose two fuzzy pharmaceutical R&D portfolio selection models with minimum borrowed capital, in which the optimal investment strategies can be generated to help investors not only maximize terminal wealth, but also minimize the cumulative borrowed capital over bankruptcy during the whole investment horizon. In the proposed models, we take into account corporate strategy in developing new products, scarcity of resources, lack of investment budget and cardinality constraint. By using the weighted sum approach, we transform the two proposed models into the corresponding single-objective models. Then, we utilize a fuzzy simulation-based ABC algorithm for solution. Our paper contributes to the existing literature on R&D in two ways. (i) This paper is the first time to investigate fuzzy R&D portfolio selection with borrowed capital. Our study gives an approach to handle R&D portfolio selection with borrowing constraint. (ii) The proposed models can provide decision maker's different aspiration levels with different investment strategies by varying the credibility levels for the practical investment requirements on R&D portfolio selection. Table 1 displays a feature comparison of the proposed R&D project portfolio models with the existing closely related researches.

Table 1. Summary of the closely related researches.

Papers	Uncertainty	Objective	Problem	Model	BC	Solution approach
Hesarsorkh et al. (2021)	fuzzy	single	R&D PP	RPP	×	GAMS/Cplex solver
Pennings and Sereno (2011)	random	single	R&D PP	COA	×	analytical solution approach
Wang and Hwang (2007)	fuzzy	single	R&D PP	COA	×	optimization technique
Hassanzadeh et al. (2012)	fuzzy	single	R&D PSS	FP	×	LINGO
Hassanzadeh et al. (2014a)	random	single	R&D PSS	RoCROP	×	robust optimization
Hassanzadeh et al. (2014b)	random	multiple	R&D PP	MBIP	×	robust optimization
Carlsson et al. (2007)	fuzzy	single	R&D PP	FMIP	×	optimization technique
Bhattacharyya et al. (2011)	fuzzy	multiple	R&D PP	FMOP	×	GA
Mohanty et al. (2005)	fuzzy	single	R&D PP	CM	×	ANP
This paper	fuzzy	multiple	R&D PP	FMOP	✓	ABC

Acronyms-BC: Borrowed capital; RPP: robust possibilistic programming; COA: compound option approach; ASA: analytical solution approach; FP:fuzzy programming; PSS: Project selection and scheduling; RoCROP: Robust CRO portfolio; MBIP: multiobjective binary integer programming; FMIP: fuzzy mixed integer programming; FMOP: fuzzy multi-objective programming; GA: genetic algorithm; CM: conceptual model; ANP: Analytical network process.

The remainder of this paper is organized as follows. In Section 2, we introduce some basic conceptions about fuzzy variables for measuring the uncertainty associated with pharmaceutical R&D projects in a fuzzy environment. In Section 3, we formulate two bi-objective fuzzy programming models for pharmaceutical R&D project portfolio selection in fuzzy environment. In Section 4, we first use the weighted sum approach to transform the proposed models into the corresponding single-objective fuzzy programming problems. Then, we use the self-dual credibility measure to deal

with the fuzzy objectives and the fuzzy constraints in the transformed models, and express investor's satisfaction degrees for the aforementioned fuzzy constraints. Thereupon, the corresponding deterministic pharmaceutical R&D project portfolio selection models can be obtained. A fuzzy simulated simulation-based artificial bee colony (ABC) algorithm is given for solution. In Section 5, we provide a numerical example to demonstrate the application of the proposed models. Finally, we conclude the paper in Section 6.

2 Basic conceptions

This section briefly reviews some concepts about fuzzy numbers, which have appeared in the literature and are necessary for our discussion in this paper.

Let ξ be a fuzzy variable defined on the possibility space $(\Theta, \mathcal{P}(\Theta), \mathrm{Pos})$ with membership function $\mu(x)$, and let u be a real number. Then, the credibility of fuzzy event $\{\xi \leq u\}$ is defined by (Liu and Liu (2002))

$$\mathrm{Cr}\{\xi \leq u\} = \frac{1}{2}(\mathrm{Pos}\{\xi \leq u\} + \mathrm{Nec}\{\xi \leq u\}), \tag{1}$$

where $\mathrm{Pos}\{\xi \leq u\} = \sup_{x \leq u} \mu(x)$ and $\mathrm{Nec}\{\xi \leq u\} = 1 - \mathrm{Pos}\{\xi > u\} = 1 - \sup_{x > u} \mu(x)$ represent the possibility and necessity measures of $\{\xi \leq u\}$, respectively. Notice that Cr is self-dual measure, i.e., $\mathrm{Cr}\{\xi \leq u\} + \mathrm{Cr}\{\xi \geq u\} = 1$. When the credibility value of a fuzzy event achieves 1, it indicates that the fuzzy event will surely happen.

In Liu (2002), we can obtain the axioms about credibility measure Cr as follows.
i) $\mathrm{Cr}\{\Theta\} = 1$.
ii) $\mathrm{Cr}\{A\} \leq \mathrm{Cr}\{B\}$ whenever $A \subset B$.
iii) $\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1$ for any event A.
iv) $\mathrm{Cr}\{\cup_i A_i\} = \sup_i \mathrm{Cr}\{A_i\}$ for any event A_i with $\sup_i \mathrm{Cr}\{A_i\} < 0.5$.
Theorem 1 (Liu (2002)). Assume that Θ is a nonempty set, $\mathcal{P}$ is the power set of Θ, and Cr is the credibility measure. Then, for any $A, B \in \mathcal{P}$, we have

$$\mathrm{Cr}\{A \cap B\} = \mathrm{Cr}\{A\} \wedge \mathrm{Cr}\{B\}, \text{if } \mathrm{Cr}\{A\} + \mathrm{Cr}\{B\} > 1. \tag{2}$$

Definition 1 (Liu and Liu (2002)). Let ξ be a fuzzy variable. The expected value of ξ is defined as

$$E(\xi) = \int_0^\infty \mathrm{Cr}\{\xi \geq t\}\mathrm{d}t - \int_{-\infty}^0 \mathrm{Cr}\{\xi \leq t\}\mathrm{d}t \tag{3}$$

provided that at least one of the two integrals is finite.
Then, the following results can be found in Liu and Liu (2002).
Theorem 2. Let ξ be fuzzy variable with finite expected values, and let λ and μ be two any given two real numbers. Then

$$E(\lambda\xi + \mu) = \lambda E(\xi) + \mu. \tag{4}$$

Theorem 3. Let ξ and η be identically distributed fuzzy variables with finite expected values, and let λ and μ be two any given nonnegative real numbers. Then

$$E(\lambda\xi + \mu\eta) = \lambda E(\xi) + \mu E(\eta). \tag{5}$$

Example 1. Let ξ be a trapezoidal fuzzy number determined by quadruplet (a, b, c, d) of crisp numbers such that $a < b < c < d$. Then, its membership function has the following form

$$\mu_\xi(x) = \begin{cases} \frac{x-a}{b-a}, \text{ if } a \leq x \leq b, \\ 1, \quad \text{if } x \in [b, c], \\ \frac{x-d}{c-d}, \text{ if } c \leq x \leq d, \\ 0, \quad \text{otherwise.} \end{cases} \tag{6}$$

By Definition 1, the expected value of the trapezoidal fuzzy variable $\xi = (a, b, c, d)$ is computed by

$$E(\xi) = \tfrac{1}{4}(a + b + c + d). \tag{7}$$

3 Mathematical formulation of pharmaceutical R&D portfolio selection

In this section, we discuss a pharmaceutical R&D portfolio selection problem with minimum borrowed capital in a fuzzy environment. Based on fuzzy set theory, we propose two pharmaceutical R&D portfolio selection models with the objectives of maximizing the terminal wealth and minimizing the cumulative borrowed capital over the whole investment periods.

3.1 Problem description and notations

Assume that a pharmaceutical company with inadequate initial capital W_0 intends to invest n candidate pharmaceutical drug R&D projects. Each candidate pharmaceutical drug R&D project contributes to a particular corporate strategy and has T_2-continuous investment periods. The whole investment horizon of each candidate pharmaceutical drug R&D project is divided into three stages including market discovery, testing and market introduction. The investment periods of market discovery and testing of each project are set as T_1 and $(T_2 - T_1)$, respectively. The time axis of each pharmaceutical drug R&D project is displayed in Fig. 1. The investment returns of these selected pharmaceutical drug projects are obtained at the end of period T_2 after market introduction. The execution of each candidate project requires the exclusive use of a number of resources (e.g., budget, human resources, etc.), while the availability of each resource type is usually limited. To avoid a possibility of bankruptcy before reaching the end of an investment horizon, we assume that the company can borrow a certain capital with borrowing rate $r_b(t)$ at the beginning of period t when

the available capital at current period is less than the corresponding actual implementation costs. In addition, we assume that the company can use the surplus capital, if possible, to invest a risk-free asset with return rate $r_f(t)$ at period $t\,(t = 1, 2, \ldots, T_2)$. Consider the fact that pharmaceutical R&D investment decision is very difficult since it is investment intensive, the required resources may be scarce and markets are ever-changing. As a result, the collected data information for portfolio management is highly uncertain and often very inaccurate. Sometimes even much of the information for project investment decision is unavailable and unreliable. Fuzzy set theory is a powerful tool to describe imprecise information in such environments. For this, we use fuzzy numbers to capture and model the uncertain pharmaceutical R&D project information. Notice that trapezoidal fuzzy variables are often used to describe the fuzzy uncertainty associated with financial markets due to their simple to estimate and easy to generalize to the LR-type forms (Carlsson et al., 2002). With the same consideration as Carlsson et al. (2002), we assume that the implement costs and the required labor to implement project $i\,(i = 1, 2, \ldots, n)$ at stage $h\,(h = 1, 2, 3)$ are characterized by trapezoidal fuzzy variables. For the sake of convenience, we introduce the following notations

$\widetilde{c}_{i,t}$ fuzzy investment costs required to pay at the beginning of period t;

$\widetilde{cp}_{i,3}$ fuzzy investment costs of pharmaceutical drug project i at stage 3;

$\widetilde{l}_{i,h}$ labor (in working months) required to implement pharmaceutical drug project i at stage h;

$\widetilde{L}_h$ labor (in working months) available to staff projects at stage h;

$\widetilde{u}_j$ maximum budget to be spent on projects contributing to strategy j;

$\widetilde{l}_j$ minimum budget to be spent on projects contributing to strategy j;

$$x_i = \begin{cases} 1, \text{ if project } i \text{ is selected for investment,} \\ 0, \text{ otherwise,} \end{cases}$$

$$s_{i,j} = \begin{cases} 1, \text{ if project } i \text{ contributes to strategy } j, \\ 0, \text{ otherwise,} \end{cases}$$

$r_f(t)$ the return rate of the risk-free asset at period $t\ (t = 1, 2, \ldots, T_2)$;

$r_b(t)$ the interest rate of borrowing capital at period $t\ (t = 1, 2, \ldots, T_2)$;

B_t the borrowed capital at the beginning of period t, where $y_t \geq 0$;

$\widetilde{d}_i$ the project returns after market introduction of project i;

W_t the available wealth at the end of period $t\ (t \in \{1, 2, \ldots, T_2\})$.

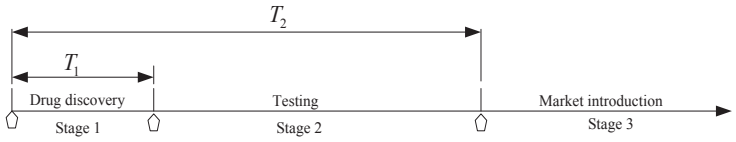

Fig. 1. Investment stages for pharmaceutical drug R&D project.

To state the model formulation, the objective functions and constraints of the proposed models are separately stated in the following two subsections.

3.2 Objectives

The total implementation costs of the n candidate projects at period t ($t = 1, 2, \ldots, T_2$) can be described by

$$C_t = \sum_{i=1}^{n} \widetilde{c}_{i,t} x_i.$$

During the process of investment, a capital shortage occurs when the available wealth is less than or equal to the required investment costs on the selected projects at current period. To avoid the occurrence of bankrupt, we assume that the borrowing capital is allowed when the company suffers capital shortage in any intermediate or the final time period. Let BE_t be the borrowing capital event at period t. Then, the borrowing capital event at period t can be described by

$$BE_t = \{W_t \leq \theta_t, W_k > \theta_k, k = 1, 2, \ldots, t-1\}, \tag{8}$$

where θ_k is the bankruptcy level at period k. It can be seen from Eq. (8) that the company does not need to borrow capital during the first $t-1$ investment periods. The available wealth obtained at the end of period k ($k \in \{1, 2, \ldots, t-1\}$) can be described as the following formula

$$W_k = (W_{k-1} - C_k)(1 + r_f(k)). \tag{9}$$

Then, we have

$$\begin{cases} W_1 &= (W_0 - C_1)(1 + r_f(1)), \\ W_2 &= (W_1 - C_2)(1 + r_f(2)), \\ &\vdots \\ W_{t-1} &= (W_{t-2} - C_{t-1})(1 + r_f(t-1)). \end{cases} \tag{10}$$

It follows from Eq. (10) that the wealth obtained at the end of period k ($k \in \{1, 2, \ldots, t-1\}$) can be rewritten as the following specific form

$$W_k = (W_0 - C_1) \prod_{i=1}^{k} (1 + r_f(i)) - \sum_{j=2}^{k} C_j \prod_{l=j}^{k} (1 + r_f(l)). \tag{11}$$

According to the analysis above, due to the shortage of capital, the company needs to borrow a certain amount of capital at each period of the following $T_2 - t$ consecutive investment periods. Then, the wealth obtained at the end of period t is calculated by

$$W_t = (W_{t-1} + B_t - C_t)(1 + r_f(t)) - B_t(1 + r_b(t)). \tag{12}$$

It follows from Eq. (12) that the formulas of the wealth obtained at the end of the following $T_2 - t$ consecutive investment periods can be represented by

$$\begin{cases} W_t &= (W_{t-1} + B_t - C_t)(1 + r_f(t)) - B_t(1 + r_b(t)), \\ W_{t+1} &= (W_t + B_{t+1} - C_{t+1})(1 + r_f(t+1)) - B_{t+1}(1 + r_b(t+1)), \\ &\vdots \\ W_{T_2} &= (W_{T_2-1} + y_{T_2} - C_{T_2})(1 + r_f(T_2)) - y_{T_2}(1 + r_b(T_2)). \end{cases} \tag{13}$$

Thus, the general form of the wealth obtained at the end of period r ($r \in \{t, t+1, \ldots, T_2\}$) is computed by

$$\begin{aligned} W_r =&(W_{t-1} + B_t - C_t)\prod_{j=t}^{r}(1+r_f(j)) - B_t(1+r_b(t))\prod_{j=t+1}^{r}(1+r_f(j)) \\ &\sum_{j=t+1}^{r} B_j(r_f(j)-r_b(j))\prod_{k=j+1}^{r}(1+r_f(k)) - \sum_{m=t+1}^{r} C_m \prod_{k=m}^{r}(1+r_f(k)). \end{aligned} \tag{14}$$

Notice that, at stage 3, the total implement costs for market introduction on the n projects is

$$CP_3 = \sum_{i=1}^{n} \widetilde{cp}_{i,3} x_i. \tag{15}$$

After finishing market introduction, the investment returns (IR) generated by the n projects is

$$IR = \sum_{i=1}^{n} \widetilde{d}_i x_i \tag{16}$$

Thus, by Eqs. (11)-(16), the terminal wealth W_T obtained at the end of the whole investment horizon can be given by

$$\begin{aligned} W_T =&W_{T_2} - CP_3 + IR \\ =&(W_0 + B_1 - C_1)\prod_{j=1}^{T_2}(1+r_f(j)) - B_1(1+r_b(1))\prod_{j=2}^{T_2}(1+r_f(j)) \\ &-\sum_{j=2}^{n} B_j(r_f(j)-r_b(j))\prod_{k=j+1}^{T_2}(1+r_f(k)) - \sum_{m=2}^{T_2}\sum_{i=1}^{n}\widetilde{c}_{i,m}x_i\prod_{k=j}^{T_2}(1+r_f(k)) \\ &-\sum_{i=1}^{n}\widetilde{cp}_{i,3}x_i + \sum_{i=1}^{n}\widetilde{d}_i x_i \end{aligned} \tag{17}$$

Assume that the first objective of the pharmaceutical company is to maximize the terminal wealth. Then, we have

$$\begin{aligned} \max W_T =&(W_0 + B_1 - C_1)\prod_{j=1}^{T_2}(1+r_f(j)) - B_1(1+r_b(1))\prod_{j=2}^{T_2}(1+r_f(j)) \\ &-\sum_{j=2}^{n} B_j(r_f(j)-r_b(j))\prod_{k=j+1}^{T_2}(1+r_f(k)) - \sum_{m=2}^{T_2}\sum_{i=1}^{n}\widetilde{c}_{i,m}x_i\prod_{k=j}^{T_2}(1+r_f(k)) \\ &-\sum_{i=1}^{n}\widetilde{cp}_{i,3}x_i + \sum_{i=1}^{n}\widetilde{d}_i x_i. \end{aligned} \tag{18}$$

Meanwhile, we assume that the pharmaceutical company is required to minimize the total borrowed capital (TBC) during the whole investment horizon. Then, we can obtain the second objective as follows:

$$\min f_{\text{TBC}} = \sum_{k=1}^{T_2+1} B_k. \tag{19}$$

3.3 Practical investment constraints

• **Borrowing constraint**. To avoid the shortage of capital, the company requires that the borrowed capital at period t, B_t, should be no less than $\max\{0, E(C_t - W_{t-1})\}$. Then, the borrowing constraint can be expressed by

$$B_t \geq \max\{0, E(C_t - W_{t-1})\},\ t \in \{1, 2, \ldots, T_2 + 1\}. \tag{20}$$

• **Budget constraint**. To keep the selected projects proceed normally, the following budget constraint about the portfolio at period t holds

$$W_{t-1} \leq \sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \leq W_{t-1} + B_t,\ t \in \{1, 2, \ldots, T_2 + 1\}. \tag{21}$$

• **Personnel constraint**. Assume that the required personnel at stage h, $\sum_{i=1}^{n} x_i \widetilde{l}_{i,h}$, should be less than or equal to the available man-power capacity $\widetilde{L}_h$. Then, the personnel constraint at stage h can be given by

$$\sum_{i=1}^{n} x_i \widetilde{l}_{i,h} \leq \widetilde{L}_h,\ h \in \{1, 2, 3\}. \tag{22}$$

• **Strategy selection constraint**. The company requires that the expenditure on projects contributing to strategy j should be restricted in the range of $[\widetilde{l}_j, \widetilde{u}_j]$. Then, the strategy selection constraint can be expressed by

$$\widetilde{l}_j \leq \sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i \leq \widetilde{u}_j,\ \forall j. \tag{23}$$

• **Cardinality constraints**. To control costs and improve portfolio performance, cardinality constraint is often incorporated into portfolio selection model. To provide pharmaceutical company with additional choices, we discuss two types of cardinality constraints as follows. (i) The pharmaceutical company requires that the maximum holding number of R&D projects in the portfolio must be not more than K. Then

$$\sum_{i=1}^{n} x_i \leq K,\ x_i \in \{0, 1\}. \tag{24}$$

(ii) The pharmaceutical company requires that the holding number of R&D projects in the portfolio must be equal to K. Then

$$\sum_{i=1}^{n} x_i = K,\ x_i \in \{0,1\}. \tag{25}$$

3.4 The proposed pharmaceutical R&D portfolio optimization models

Notice that, in real life, pharmaceutical R&D is often very expensive and requires substantial capital expenditure. Pharmaceutical R&D project portfolio is a very complex decision-making process since it is affected by many critical factors such as funds and resources. During the process of pharmaceutical R&D project portfolio, due to the limitation of funds and resources, it is very important for pharmaceutical company to determine a suitable number of R&D projects in a portfolio. To provide decision makers more choices, we discuss two pharmaceutical R&D project portfolio selection problems with different cardinality constraints in this section. By using fuzzy set theory, we propose two pharmaceutical R&D project portfolio selection models with minimum borrowed capital. In the two models, we assume that the pharmaceutical company intends to seek optimal investment strategies with the objectives of maximizing terminal wealth and minimizing the cumulative borrowing capital over the whole investment horizon.

(1) Assume that the pharmaceutical company considers decision criteria including budget constraint (21), personnel constraint (22) and strategy selection constraint (23). Moreover, it requires that the maximum holding number of projects in the portfolio must not exceed K. Following the idea, we formulate a fuzzy programming model (P_1) for pharmaceutical R&D portfolio selection as follows

$$(P_1)\begin{cases} \max & W_T = (W_0 + B_1 - C_1)\prod_{j=1}^{T_2}(1+r_f(j)) - B_1(1+r_b(1))\prod_{j=2}^{T_2}(1+r_f(j)) \\ & \sum_{j=2}^{n} B_j(r_f(j)-r_b(j))\prod_{k=j+1}^{T_2}(1+r_f(k)) - \sum_{m=2}^{T_2}\sum_{i=1}^{n}\widetilde{c}_{i,m}\prod_{k=j}^{T_2}(1+r_f(k)) \\ & -\sum_{i=1}^{n}\widetilde{cp}_{i,3}x_i + \sum_{i=1}^{n}\widetilde{d}_i x_i \\ \min & f_{\text{TBC}} = \sum_{k=1}^{T_2+1} B_k \\ s.t. & \text{Constraints } (20) \sim (24). \end{cases}$$

(2) In this model, we assume that the company uses constraint (25) rather than constraint (24) to limit the holding number of projects in the portfolio. Namely, the company requires that the holding number of projects in the portfolio is equal to K. In addition, the company takes the two decision objectives and all other investment constraints in the model (P_1) into consideration. Then, we formulate the following

fuzzy pharmaceutical R&D portfolio selection model (P_2):

$$(P_2)\begin{cases}\max & W_T \\ \min & f_{\text{TBC}} = \sum\limits_{k=1}^{T_2+1} B_k \\ s.t. & \text{Constraints } (20), (21), (22), (23) \text{ and } (25).\end{cases}$$

Based on the analysis above, we can find that the difference between the models (P_1) and (P_2) is the cardinality constraint.

4 Solution approach

Notice that the proposed two models are bi-objective programming problems with fuzzy coefficients in both the objective functions and constraints. In this section, we give a fuzzy simulation-based artificial bee colony algorithm for solution. Here, we take the model (P_1) as an example to introduce the solution approach.

The model (P_1) can be equivalently rewritten into a vector optimization problem with minimization objectives as follows:

$$(P_1)' \begin{cases}\min & \{-W_T, f_{\text{TBC}}\} \\ s.t. & \text{Constraints } (20) \sim (24).\end{cases}$$

Notice that the model $(P_1)'$ is a bi-objective programming problem. To handle this kind of problem, the most widely used methods are the weighted sum approach in Marler and Arora (2010) and the ϵ-constraint method in Chankong and Haimes (1983). The main defect of the ϵ-constraint method is that it is very sensitive to the number of objectives (Copado-Méndez et al., 2016). As mentioned by Marler and Arora (2010), the weighted sum approach can not only generate multiple solutions by varying the weights, but also provide a single solution that describes preferences presumably associated with the selection of a group of weights. For this, we use the weighted sum approach to transform the model $(P_1)'$ into the following single objective programming problem

$$(P_1)'' \begin{cases}\min & \omega f_{\text{TBC}} - (1-\omega) W_T \\ s.t. & \text{Constraints } (20) \sim (24),\end{cases}$$

where $\omega \in [0, 1]$ is the relative preference weight of the pharmaceutical company with respect to f_{TBC}. $\omega = 0$ indicates that the company only focuses on the objective W_T. $\omega = 0.5$ indicates that the company shows indifferent attitudes to both the objectives W_T and f_{TBC}. $\omega = 1$ indicates that the company only considers the objective f_{TBC}.

4.1 Deterministic pharmaceutical R&D project portfolio optimization model

Since the model $(P_1)''$ is a fuzzy programming problem with fuzzy coefficients in both objective function and investment constraints, a commonly used technique to

transform it into the corresponding deterministic programming problem. In this paper, we use the self-dual credibility measure as the basic technique tool to transform these fuzzy coefficients in the objective function and the investment constraints of the model $(P_1)''$ into the deterministic forms and express investor's aspiration levels for the satisfaction of the fuzzy coefficient constraints.

4.1.1 Treatment of the fuzzy objective function

The implementation costs on the candidate pharmaceutical R&D projects at each stage and the obtained revenues are all fuzzy numbers. Derived from the extension principle in Zadeh (1965), the scalar multiplication and additive operations on several fuzzy numbers is also a fuzzy number. It follows from Eq. (13) that the terminal wealth obtained at the end of stage 3 is also a fuzzy number. By Definition 1, Theorems 2 and 3, the expected value of the terminal wealth is

$$\begin{aligned}\max E(W_T) =&(W_0 + B_1 - E(C_1)) \prod_{j=1}^{T_2}(1 + r_f(j)) - B_1(1 + r_b(1)) \prod_{j=2}^{T_2}(1 + r_f(j)) \\ &\sum_{j=2}^{n} B_j(r_f(j) - r_b(j)) \prod_{k=j+1}^{T_2} (1 + r_f(k)) - \sum_{m=2}^{T_2} \sum_{i=1}^{n} E(\widetilde{c}_{i,m}) \prod_{k=j}^{T_2}(1 + r_f(k)) \\ &- \sum_{i=1}^{n} E(\widetilde{cp}_{i,3})x_i + \sum_{i=1}^{n} E(\widetilde{d}_i)x_i. \end{aligned} \tag{26}$$

Thus, the objective function of the model (P_1'') can be expressed as the following deterministic form

$$\min \ \ \omega f_{\text{TBC}} - (1 - \omega)E(W_T). \tag{27}$$

4.1.2 Treatment of the fuzzy constraints

Due to the exist of fuzzy coefficients in investment constraints, it leads these constraints to be not clear at all. So, the above-mentioned investment constraints with fuzzy coefficients may be not strictly held. In this case, it is necessary to take into account a certain deviation and flexibility on the holding of these fuzzy constraints. Similar to Liu and Iwamura (1998), we use the self-dual credibility measure as a tool to handle these fuzzy constraints and describe the aspiration levels of the pharmaceutical R&D company for the satisfaction of these fuzzy constraints in this paper. Assume that the investment constraints (21), (22) and (23) hold with at least credibility levels λ, θ and η, respectively. Then, the three fuzzy investment constraints can

be represented by the chances of the credibility as follows

$$\mathrm{Cr}(W_{t-1} \leq \sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \leq W_{t-1} + B_t) \geq \lambda,\ t \in \{1, 2, 3, \ldots, T_2 + 1\}, \tag{28}$$

$$\mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{l}_{i,h} \leq \widetilde{L}_h) \geq \theta,\ h \in \{1, 2, 3\}, \tag{29}$$

$$\mathrm{Cr}(\widetilde{l}_j \leq \sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i \leq \widetilde{u}_j) \geq \eta,\ \forall j. \tag{30}$$

For the sake of description, we set $A = \sum_{i=1}^{n} x_i \widetilde{c}_{i,t} > W_{t-1} + B_t$ and $B = \sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \geq W_{t-1}$. Then, we have $A^c = \sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \leq W_{t-1} + B_t$ and $A \subset B$. It follows from ii) that $\mathrm{Cr}\{A\} \leq \mathrm{Cr}\{B\}$. That is,

$$\mathrm{Cr}\{\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} > W_{t-1} + B_t\} \leq \mathrm{Cr}\{\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \geq W_{t-1}\}. \tag{31}$$

Therefore, by iii), we get

$$\mathrm{Cr}\{\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} > W_{t-1} + B_t\} + \mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \leq W_{t-1} + B_t) = 1. \tag{32}$$

According to Eqs. (31) and (32), we see that

$$\mathrm{Cr}\{\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \geq W_{t-1}\} + \mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \leq W_{t-1} + B_t) > 1 \tag{33}$$

By Theorem 1, Eq (28) can be equivalently rewritten into the following form

$$\mathrm{Cr}(W_{t-1} \leq \sum_{i=1}^{n} x_i \widetilde{c}_{i,t}) \wedge \mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \leq W_{t-1} + B_t) \geq \lambda. \tag{34}$$

Similarly, Eq. (30) can be equivalently represented by

$$\mathrm{Cr}(\widetilde{l}_j \leq \sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i) \wedge \mathrm{Cr}(\sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i \leq \widetilde{u}_j) \geq \eta,\ \forall j. \tag{35}$$

Then, the model $(P_1)''$ can be rewritten into the following credibility-based chance-constrained programming problem

$$(P_1)''' \begin{cases} \min & \omega \sum_{t=1}^{T_2+1} B_t - (1-\omega)E(W_T) \\ s.t. & \mathrm{Cr}(W_{t-1} \le \sum_{i=1}^{n} x_i \widetilde{c}_{i,t}) \wedge \mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \le W_{t-1} + B_t) \ge \lambda, \\ & \mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{l}_{i,h} \le \widetilde{L}_h) \ge \theta,\ h \in \{1,2,3\}, \\ & \mathrm{Cr}(\widetilde{l}_j \le \sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i) \wedge \mathrm{Cr}(\sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i \le \widetilde{u}_j) \ge \eta,\ \forall j, \\ & \sum_{i=1}^{n} x_i \le K,\ x_i \in \{0,1\}, \\ & y_t \ge \max\{0, E(C_t - W_{t-1})\},\ t \in \{1,2,\ldots,T_2+1\}. \end{cases}$$

Notice that the model $(P_1)'''$ has some fuzzy credibility-based chance constraints. As mentioned by Liu and Iwamura (1998) and Liu (2002), in fuzzy environment, the basic technique for handling this kind of problem is to transform it into the corresponding crisp equivalent under the predetermined confidence level. However, this approach is often very hard and only successful for several special cases. For this, we apply the fuzzy simulation technique in Liu and Iwamura (1998) to calculate the objective value and the credibility values of fuzzy variables in the model $(P_1)'''$.

4.1.3 Fuzzy simulation technique

Before solving the model $(P_1)'''$, we need to calculate the uncertain values as follows

Expected Values:$E(W_T)$ and $E(C_t - W_{t-1})$,

$$U_1(x) : \mathrm{Cr}(W_{t-1} - \sum_{i=1}^{n} x_i \widetilde{c}_{i,t} \le 0) \wedge \mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{c}_{i,t} - (W_{t-1} + B_t) \le 0) \ge \lambda,\ t \in \{1,2,\ldots,T_2+1\},$$

$$U_2(x) : \mathrm{Cr}(\sum_{i=1}^{n} x_i \widetilde{l}_{i,h} - \widetilde{L}_h \le 0) \ge \theta,\ h \in \{1,2,3\},$$

$$U_3(x) : \mathrm{Cr}(\widetilde{l}_j - \sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i \le 0) \wedge \mathrm{Cr}(\sum_{i=1}^{n} \sum_{t=1}^{T_2} s_{i,j} \widetilde{c}_{i,t} x_i - \widetilde{u}_j \le 0) \ge \eta,\ \forall j.$$

In this paper, fuzzy simulation technique introduced in Liu (2002) is used to compute the credibility measures for the aforementioned uncertain values.

• *Fuzzy simulation for expected value*

By Definition 1, the expected values of W_T and $C_t - W_{t-1}$ can be, respectively, calculated by

$$E(W_T) = \int_0^\infty \mathrm{Cr}\{W_T \geq \gamma\}\mathrm{d}\gamma - \int_{-\infty}^0 \mathrm{Cr}\{W_T \leq \gamma\}\mathrm{d}\gamma, \tag{36}$$

$$E(C_t - W_{t-1}) = \int_0^\infty \mathrm{Cr}\{C_t - W_{t-1} \geq \gamma\}\mathrm{d}\gamma - \int_{-\infty}^0 \mathrm{Cr}\{C_t - W_{t-1} \leq \gamma\}\mathrm{d}\gamma. \tag{37}$$

•*Fuzzy simulation for $U_1(x), U_2(x)$ and $U_3(x)$*

The general form of $U_1(x), U_2(x)$ and $U_3(x)$ can be expressed by $U(x) = \mathrm{Cr}\{\xi | f(x,\xi) \leq 0\}$, where $x = (x_1, x_2, \ldots, x_n)$ is a decision variable vector and $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ is a fuzzy variable vector. To calculate the value of $U(x)$, we randomly produce ξ_m and set $\nu_m = \mu(\theta_m)$ for $m = 1, 2, \ldots, N$, where μ is the membership function of ξ. Then, the credibility of the uncertain function $U(x)$ is computed by

$$U(x) = \tfrac{1}{2}\big(\max_{1\leq m\leq N}\{\nu_m | f(x,\xi_m) \geq 0\} + \min_{1\leq m\leq N}\{1 - \nu_m | f(x,\xi_m) < 0\}\big). \tag{38}$$

The fuzzy simulation process for computing $U(x)$ is summarized as follows:
Step 1. Generate a random number θ_m from the credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$, set $\nu_m = (2\mathrm{Cr}\{\theta_m\}) \wedge 1$ and produce $\xi_m = \xi(\theta_m), m = 1, 2, \ldots, N$, respectively.
Step 2. Return $U(x)$ via the estimation formula as shown in Eq. (38).

For the uncertain function $U_1(x)$, we set $f_1(x,\xi) = W_{t-1} - \sum_{i=1}^n x_i\widetilde{c}_{i,t}$ and $f_2(x,\xi) = \sum_{i=1}^n x_i\widetilde{c}_{i,t} - (W_{t-1} + B_t)$. The credibility $U_1(x)$ is estimated by

$$\begin{aligned} U_1(x) = {} & \tfrac{1}{2}\big(\max_{1\leq m\leq N}\{\nu_k | f_1(x,\xi_m) \geq 0\} + \min_{1\leq m\leq N}\{1 - \nu_m | f_1(x,\xi_m) < 0\}\big) \\ & \wedge \tfrac{1}{2}\big(\max_{1\leq m\leq N}\{\nu_m | f_2(x,\xi_m) \geq 0\} + \min_{1\leq m\leq N}\{1 - \nu_m | f_2(x,\xi_m) < 0\}\big). \end{aligned}$$

To compute $U_2(x)$, we set $f(x,\xi) = \sum_{i=1}^n x_i\widetilde{l}_{i,k} - \widetilde{L}_k$. Then, its credibility can be estimated by Eq. (38). For $U_3(x)$, we set $f_3(x,\xi) = \widetilde{l}_j - \sum_{i=1}^n\sum_{t=1}^{T_2} s_{i,j}\widetilde{c}_{i,t}x_i$ and $f_4(x,\xi) = \sum_{i=1}^n\sum_{t=1}^{T_2} s_{i,j}\widetilde{c}_{i,t}x_i - \widetilde{u}_j$. Then, the credibility of $U_3(x)$ is estimated by

$$\begin{aligned} U_3(x) = {} & \tfrac{1}{2}\big(\max_{1\leq m\leq N}\{\nu_k | f_3(x,\xi_m) \geq 0\} + \min_{1\leq m\leq N}\{1 - \nu_m | f_3(x,\xi_m) < 0\}\big) \\ & \wedge \tfrac{1}{4}\big(\max_{1\leq m\leq N}\{\nu_k | f_4(x,\xi_m) \geq 0\} + \min_{1\leq m\leq N}\{1 - \nu_m | f_4(x,\xi_m) < 0\}\big). \end{aligned}$$

• *Artificial bee colony algorithm*

Notice that the proposed models with fuzzy variables in both the objective functions and constraints have cardinality constraints. Thus, they are NP-hard problems. Due to their mathematical complexity, traditional optimization methods often fail

to solve them. For solution, we must look for meta-heuristic algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony Algorithm (ABC). Compared with other methods, the ABC algorithm is a relatively new meta-heuristic algorithm with fewer control parameters, a simpler structure, and more convincing performance (Karaboga and Akay, 2009). For this, we employ a modified artificial bee colony algorithm (ABC) in Karaboga and Akay (2011) to solve our models. In the following, we take the model $(P_1)''$ as an example to introduce the solution algorithm.

ABC algorithm is a meta-heuristic technique proposed by Karaboga (2005), which is inspired by simulating the foraging behavior of honey bees. In ABC, the food source of bees is viewed as a potential solution of the optimization problem and the nectar amount of the food source denotes the quality (fitness value) of the corresponding solution. ABC consists of three kinds of bees (i.e., employed bees, onlooker bees and scout bees). The population sizes of both employed bees and onlooker bees are the same, which is equal to the number of solutions in the population. Among them, different kinds of bees often have different tasks. The main task of employed bees is to look for food source around their hive and share the information about food source with onlooker bees by waggle dance. The main task of onlooker bees is to explore some better food sources for further search. The main task of scout bees is to perform a random search to find new food sources in the search space. When the nectar amount of a food source is low or exhausted, the food source will be abandoned by employed bees and these employed bees will become scout bees to search new food sources. ABC algorithm has been proved to be a promising performance on various complex optimization problems and received more and more attention during the past few years. For a detailed discussion, one can refer to Bajer and Zorić (2019), Li and Yang (2016), Xiang et al. (2018), Karaboga and Akay (2011) and Bansal et al. (2018).

As mentioned by Karaboga and Akay (2011), the modified ABC algorithm is effective for constrained optimization problems and shows better performance than that of state-of-the-art algorithms. For this, we use the modified ABC algorithm in Karaboga and Akay (2011) to solve the model $(P_1)''$ in this paper. In the following, let us introduce its initialization, fitness function, constraint-handling, employed bee phase, onlooker bee phase and scout bee phase.

Initialization. Randomly generate a swarm of food sources (potential solutions) with population size pop_size. For a numerical optimization problem, each food source consists of a D-dimensional parameter vector. Here, D is equal to the number of variables in the optimization problem. Similar to the other evolutionary algorithms, ABC generates an initial population of food sources randomly. Let $\chi = (x, B)$ be a solution of the model $(P_1)''$, where χ has two vectors including a binary vector $x = (x_1, x_2, \ldots, x_n)$ and a real-valued vector $B = (B_1, B_2, \ldots, B_{T_2+1})$. The binary vector x controls the holding number of projects in the portfolio. The real-valued vector B controls the borrowed capital over the whole investment horizon. Then, a solution χ of the model $(P_1)''$ can be encoded as a food source with the following form

$$x_i = x_i^{\min} + (x_i^{\max} - x_i^{\min}) \cdot rand_1(0,1),\ i = 1, 2, \ldots, n, \tag{39}$$

$$B_t = B_t^{\min} + (B_t^{\max} - B_t^{\min}) \cdot rand_2(0,1),\ t = 1, 2, \ldots, y_{T_2+1}, \tag{40}$$

where $rand_1(0,1)$ and $rand_2(0,1)$ are two uniformly distributed random numbers in $[0,1]$; $x_i^{\min}$ and $x_i^{\max}$ are the lower and upper bounds of x_i, respectively; $B_t^{\min}$ and $B_t^{\max}$ are the lower and upper bounds of B_t, respectively. Repeat the above process pop_size times, we can obtain a population with pop_size initialized individuals.

Fitness function. Notice that the model $(P_1)''$ is a minimization problem. After initialization, the fitness value of each solution is calculated as follows

$$fit_i = \begin{cases} \frac{1}{(1+f_i)}, & \text{if } f_i \geq 0, \\ 1 + |f_i|, & \text{otherwise} \end{cases} \tag{41}$$

where fit_i and f_i represent the fitness value and objective function value of the solution χ_i, respectively.

Constraint-handling. To satisfy Eq. (24), we sort the values of $x_1, x_2, \ldots, x_n$ in descending order. After that, we keep the first K values of x_i and set all other x_i to zero. The violation value of the ith solution, V_i, is computed by

$$V_i = \sum \max\{g_p(\chi_i), 0\}, \tag{42}$$

where $g_p(\chi_i) \leq 0$ denote the pth inequality constraint of the model $(P_1)''$.

Employed bees phase. In this phase, each employed bee searches around a distinct food source position and generates a new candidate food source position $V = (V_1, V_2, \ldots, V_D)$ with the following form

$$V_i^j = \begin{cases} \chi_i^j + rand(-1,1)(\chi_i^j - \chi_i^k), \text{ if } rand_i < MR, \\ \chi_i^j, \qquad\qquad\qquad\qquad \text{otherwise.} \end{cases} \tag{43}$$

where $i \in \{1, 2, \ldots, pop\_size\}$ and $j \in \{1, 2, \ldots, D\}$; k is a randomly chosen index, $k \in \{1, 2, \ldots, pop\_size\}$, and $k \neq i$; $rand(-1,1)$ is a uniformly distributed random real number in $[-1,1]$; $rand_i$ is a random number in the interval $[0,1]$ and MR is a control parameter of ABC algorithm in the interval $[0,1]$, which controls the number of parameters to be modified. If no parameter is changed, we need to change one random parameter of the solution χ_i by $\chi_i^j + rand(-1,1)(\chi_i^j - \chi_i^k)$. After that, the fitness value of the new position V_i is computed. If the fitness value of new position V_i is better than the one of χ_i, we select the solution vector V_i, otherwise keep χ_i.

After producing a new food source (solution), ABC algorithm makes a selection. In this algorithm, we apply the rules in Deb (2000) to solve constrained optimization problems. Based on Deb's rules, for any given two solution χ_i and χ_j, we replace χ_j by χ_i with the following criteria:
i) When χ_i is feasible and χ_j is infeasible;
ii) When both χ_i and χ_j are feasible with $fit_i < fit_j$;
iii) When both χ_i and χ_j are infeasible with $V_i < V_j$.

Onlooker bees phase. After all employed bees complete the search process, the onlooker bees will collect the information from all of the employed bees and choose

good food sources to gather honey depending on the probability value. In this algorithm, the probability of the food source i chosen by an onlooker bee is denoted by p_i with the following form

$$p_i = \begin{cases} 0.5 \times (1 + \frac{fit_i}{\sum_{j=1}^{pop\_size} fit_j}), \text{ if solution is feasible,} \\ 0.5 \times (1 - \frac{V_i}{\sum_{j=1}^{pop\_size} V_j}), \quad \text{otherwise.} \end{cases} \tag{44}$$

Notice that, in this algorithm, the infeasible solutions as well as the feasible ones are allowed to populate in the colony. The feasible solutions are selected probabilistically proportional to their fitness values, and the infeasible solutions are selected inversely proportional to their violation values. The probabilities of the infeasible solutions are between 0 and 0.5, while the probabilities of the feasible ones are between 0.5 and 1. At this phase, onlooker bee generates a modified position of the food source as shown in Eq. (43) based on the probability value and checks its nectar amount. If a parameter value produced by this operation exceeds its predetermined boundaries, the parameter can be set as an acceptable value.

Scout bees phase. Define *limit* and $fail_i$ as the predetermined number of trials and the iteration number of failing to explore better solutions for individual i, respectively. In this phase, once a food source cannot be improved after a predetermined number of iterations, i.e., $fail_i > limit$. Then, this food source will be abandoned from the population. By using Eqs. (39-40), to keep the population diversity, the abandoned food source will be updated by a new random food source discovered by the scouts. In this procedure, another control parameter called scout production period (*SPP*) is also used to enhance the convergence capability of the algorithm for constrained optimization problems. At each *SPP*, when an abandoned food source exceeds *limit*, a scout production process will be performed to provide a diversity mechanism by allowing new and probably infeasible food sources in the population.

The procedures of the algorithm are described as follows.

Step 1. Input the algorithm parameters including colony size CS, modification rate MR, maximum cycle number MCN, predetermined number of cycles $limit$ and scout production period SPP;

Step 2. Initialize food sources by Eqs. (39) and (40);

Step 3. Calculate the fitness values of all food sources by Eq. (41) and memorize the best food source achieved so far;

Step 4. Perform employed bees phase by Eq. (43);

Step 5. Calculate the fitness value of the new food source and select the better food source position based on Deb's rules;

Step 6. Calculate probabilities for onlookers by Eq. (44) and perform onlooker bees phase;

Step 7. Perform scout bees phase and replace the abandoned food sources by Eqs. (39) and (40);

Step 8. Return to Step 3 until the termination criteria is met.

Step 9. Report the best solution found so far as the optimum solution for the problem.

5 Numerical example

In this section, we use a numerical example to illustrate the application of the proposed pharmaceutical R&D project portfolio models.

In this example, we consider a pharmaceutical R&D project portfolio selection problem for a pharmaceutical company with 20 candidate pharmaceutical R&D projects (P1–P20), which is similar to Hassanzadeh et al. (2012) and Rogers et al. (2002). Assume that each candidate pharmaceutical R&D project has three investment stages including drug discovery, testing and market introduction. The drug discovery phase and the testing phase of all projects are assumed to take three years and seven years, respectively. The investment costs and revenues are discounted to the beginning of the planning horizon. Tables 2 and 3 show the development costs and the required R&D human resources represented by fuzzy numbers for the three investment stages. The initial budget for the three investment stages is set to be 150 (in millions) , i.e., $W_0 = 150$. The preset capacities of staff projects for the three stages are (in working months) set to be fuzzy numbers (374.5, 374.5, 0, 50), (1964.9, 1964.9, 0, 250), and (1319.5, 1319.5, 0, 160), respectively. Table 2 lists the fuzzy costs and the revenues of the 20 candidate R&D projects at each stage discounted to the beginning of stage 1 with interest rates of 4% and 15%, respectively. Namely, $r_c = 4\%$ and $r_d = 15\%$. The fuzzy costs of the 20 candidate R&D projects at each stage are required to be averagely apportioned among the corresponding investment periods of the current stage. Moreover, the investment strategies of the 20 candidate R&D projects are classified into three strategic types including new drug (S1={P13, P14, P17, P18, P19, P20}), derivative of existing drug (S2 ={P5, P6, P8, P9, P10, P15, P16}) and incremental improvement to existing drugs (S3 = {P1, P2, P3, P4, P7, P11, P12}). The target balances across three R&D strategies are 40-70%, 20-40%, and 10-30%, respectively. The risk-free rate and the borrowing rate over the whole investment horizon are set as 2.5% and 4.5%, respectively.

Table 2. Fuzzy present value of costs/revenues for 20 candidate projects (in millions).

Project no.	Stage 1 cost	Stage 2 cost	Stage 3 cost	Stage 3 revenue
P1	(2,2.2,0.3,0.3)	(33.7470,37.1217,5.0621,5.0621)	(44.4060,48.8466,6.6609,6.6609)	(74.01,81.4110,11.1015,11.1015)
P 2	(3,3.3,0.45,0.45)	(56.2450,61.8695,8.4368,8.4368)	(66.6090,73.2699,9.9914,9.9914)	(148.02,162.8220,22.2030,22.2030)
P 3	(10,11,1.5,1.5)	(84.3675,92.8042,12.6551,12.6551)	(148.0200,162.8220,22.2030,22.2030)	(296.04,325.6440,44.4060,44.4060)
P 4	(5,5.5,0.75,0.75)	(73.1185,80.4304,10.9678,10.9678)	(251.6340,276.7974,37.7451,37.7451)	(296.04,325.6440,44.4060,44.4060)
P 5	(20,22,3,3)	(95.6165,105.1782,14.3425,14.3425)	(296.0400,325.6440,44.4060,44.4060)	(888.12,976.9320,133.2180,133.2180)
P 6	(15,16.5,2.25,2.25)	(44.9960,49.4956,6.7494,6.7494)	(66.6090,73.2699,9.9914,9.9914)	(148.02,162.8220,22.2030,22.2030)
P 7	(7,7.7,1.05,1.05)	(39.3715,43.3087,5.9057,5.9057)	(44.4060,48.8466,6.6609,6.6609)	(118.416,130.2576,17.7624,17.7624)
P 8	(5,5.5,0.75,0.75)	(61.8695,68.0564,9.2804,9.2804)	(74.0100,81.4110,11.1015,11.1015)	(148.02,162.8220,22.2030,22.2030)
P 9	(10,11,1.5,1.5)	(84.3675,92.8042,12.6551,12.6551)	(118.4160,130.2576,17.7624,17.7624)	(266.436,293.0796,39.9654,39.9654)
P 10	(18,19.8,2.7,2.7)	(95.6165,105.1782,14.3425,14.3425)	(177.6240,195.3864,26.6436,26.6436)	(562.4760,618.7236,84.3714,84.3714)
P 11	(5,5.5,0.75,0.75)	(39.3715,43.3087,5.9057,5.9057)	(44.4060,48.8466,6.6609,6.6609)	(118.416,130.2576, 17.7624, 17.7624)
P 12	(7,7.7,1.05,1.05)	(44.9960,49.4956,6.7494,6.7494)	(88.8120,97.6932,13.3218,13.3218)	(148.02,162.8220,22.2030,22.2030)
P 13	(15,16.5,2.25,2.25)	(106.8655,117.5520,16.0298,16.0298)	(266.4360,293.0796,39.9654,39.9654)	(59.208,65.1288,8.8812,8.8812)
P 14	(35,38.5,5.25,5.25)	(134.9880,148.4868,20.2482,20.2482)	(414.4560,455.9016,62.1684,62.1684)	(1036.14,1139.754,155.421,155.421)
P15	(25,27.5,3.75,3.75)	(78.7430,86.6173,11.8115,11.8115)	(148.0200,162.8220,22.2030,22.2030)	(740.10,814.1100,111.0150,111.0150)
P 16	(15,16.5,2.25,2.25)	(106.8655,117.5520,16.0298,16.0298)	(222.0300,244.2330,33.3045,33.3045)	(444.06,488.4660,66.6090,66.6090)
P17	(17,18.7,2.55,2.55)	(89.9920,98.9912,13.4988,13.4988)	(266.4360,293.0796,39.9654,39.9654)	(518.07,569.8770,77.7105,77.7105)
P 18	(20,22,3,3)	(101.2410,111.3651,15.1862,15.1862)	(325.6440,358.2084,48.8466,48.8466)	(814.11,895.5210,122.1165,122.1165)
P 19	(35,38.5,5.25,5.25)	(134.9880,148.4868,20.2482,20.2482)	(370.05,407.055,55.5075,55.5075)	(1184.16,1302.576,177.624,177.624)
P20	(50,55,7.5,7.5)	(146.2370,160.8607,21.9355,21.9355)	(518.0700,569.8770,77.7105,77.7105)	(1702.23,1872.453,255.3345,255.3345)

Table 3. Required labor represented by fuzzy numbers for 20 candidate projects.

Project no.	fuzzy development resource (in working months)		
	Stage 1	Stage 2	Stage 3
P1	(6,6,0.6,0.6)	(72,72,7.2,7.2)	(50,50,5,5)
P2	(12,12,1.2,1.2)	(80,80,8,8)	(48,48,4.8,4.8)
P3	(24,24,2.4,2.4)	(95,95,9.5,9.5)	(70,70,7,7)
P4	(12,12,1.2,1.2)	(100,100,10,10)	(70,70,7,7)
P5	(32,32,3.2,3.2)	(120,120,12,12)	(80,80, 8,8)
P6	(26,26,2.6,2.6)	(105,105,10.5,10.5)	(75,75,7.5,7.5)
P7	(20,20,2,2)	(85,85,8.5,8.5)	(52,52,5.2,5.2)
P8	(12,12,1.2,1.2)	(110,110,11,11)	(75,75,7.5,7.5)
P9	(24,24,2.4,2.4)	(150,150,15,15)	(90,90,9,9)
P10	(30,30,3,3)	(155,155,15.5,15.5)	(100,100,10,10)
P11	(14,14,1.4,1.4)	(90,90,9,9)	(60,60,6,6)
P12	(15,15,1.5,1.5)	(75,75,7.5,7.5)	(70,70,7,7)
P13	(30,30,3,3)	(180,180,18,18)	(120,120,12,12)
P14	(45,45,4.5,4.5)	(200,200,20,20)	(130,130,13,13)
P15	(40,40,4,4)	(160,160,16,16)	(110,110,11,11)
P16	(35,35,3.5,3.5)	(190,190,19,19)	(125,125,12.5,12.5)
P17	(36,36,3.6,3.6)	(190,190,19,19)	(120,120,12,12)
P18	(38,38,3.8,3.8)	(200,200,20,20)	(130,130,13,13)
P19	(36,36,3.6,3.6)	(220,220,22,22)	(150,150,15,15)
P20	(48,48,4.8,4.8)	(230,230,23,23)	(160,160,16,16)

We apply the two proposed models to this example and use the modified ABC algorithm in Karaboga and Akay (2011) for solution. In the two models, the credibility levels λ, θ and η are set as 0.8, 0.8 and 0.8, respectively. We set the relative preference weight ω to be 0, 0.5 and 1. The parameter values of the solution algorithm are set as follows. The value of modification rate (MR) is set as 0.8. The value of colony size is set as 40, i.e., $CS = 40$. The maximum cycle number is set as 2000, i.e., $MCN = 2000$. The value of the predetermined number of trials $limit$ is equal to $0.5 \times CS \times D$, where D is the dimension of the model and CS is the number of solutions in the population. The value of scout production period SPP is also equal to $0.5 \times CS \times D$. We run the aforementioned algorithm with 2000 generations on each model. The corresponding comparative computational results of the two models about the selected projects under different strategic types, project size, the amount of the borrowed capital at period t (B_t) and terminal wealth are displayed in Table 4.

Table 4. Comparative computational results obtained by the models (P_1) and (P_2).

Decision index	Model (P_1)			Model (P_2)		
	$\omega = 0$	$\omega = 0.5$	$\omega = 1$	$\omega = 0$	$\omega = 0.5$	$\omega = 1$
Strategy 1	{P18, P19, P20}	{P17, P18, P19, P20}	{P19}	{P14, P19, P20}	{P13, P14, P18, P20}	{P18}
Strategy 2	{P5, P10, P15}	{P5, P10, P15}	{P10}	{P5, P9, P10}	{P5, P10, P16 }	{P6, P8, P9}
Strategy 3	{P2, P3, P4, P12}	{P3, P4}	{P3, P11}	{P2, P4, P11, P12}	{P3, P7, P11}	{P1, P2, P3, P4, P11, P12}
Project size	10	9	4	10	10	10
B_1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
B_2	52.553411	59.820908	0.000000	48.699496	32.147673	0.000000
B_3	131.54371	140.30464	0.000000	125.43089	100.46098	0.000000
B_4	292.34163	299.62821	45.345607	284.92393	100.46098	104.77443
B_5	460.45629	466.33176	103.65873	451.64195	401.85462	213.25862
B_6	636.01929	640.75732	172.29099	625.9423	564.02363	329.63338
B_7	819.61172	822.36626	235.67249	808.15632	732.50916	454.40428
B_8	1011.4152	1011.9394	312.60676	998.37765	909.51039	580.06282
B_9	1211.7648	1210.9112	380.42525	1197.1822	1093.0892	712.24144
B_{10}	1421.4287	1418.1828	458.82989	1405.0462	1285.6154	848.29299
B_{11}	6636.0945	4458.8024	1361.5664	6865.124	4082.4702	2352.2456
$E(W_T)$	43943.1422	16158.6654	3070.3961	41740.0501	13518.6544	4989.3997

It can be seen from Table 4 that the optimal investment strategies obtained by the two models under different ω are different. Let us take $\omega = 0.5$ as an example to introduce the investment strategies of the two models. If the pharmaceutical

company makes a portfolio decision based on the model (P_1), then it should follow the investment strategy listed in Column 2 of Table 4 to adjust its wealth during the whole investment horizon. In this case, the pharmaceutical company should select 9 projects (including projects P4, P5, P10, P11, P12, P17, P18, P19 and P20) to construct a portfolio. P17, P18, P19 and P20 should be selected for strategy 1. P5, P10 and P15 should be selected for strategy 2. P3 and P4 should be selected for strategy 3. The total borrowed capital over the whole investment horizon is 10529.044898 million. The terminal wealth obtained at the end of investment horizon is 16158.6654 million. However, if the pharmaceutical company makes a portfolio decision by using the model (P_2), then it should follow the investment strategy listed in Column 3 of Table 4 to adjust its wealth at the beginning of each investment period. In this case, 10 projects are selected to construct a portfolio. For strategy 1, P17, P18, P19 and P20 should be selected. For strategy 2, P5 and P10 should be selected. For strategy 3, P4, P11 and P12 should be selected. During the process of investment, the total borrowed capital is 9449.218183 millon. At the end of stage 3, the obtained terminal wealth is 113518.6544 million. Thus, we can conclude that, in term of terminal wealth, the model (P_1) performs better than the model (P_2).

In the following, we perform further numerical experiment to investigate the effects of the decision maker's different aspiration levels for the practical investment requirements including budget constraint, personnel constraint and personnel constraint on project portfolio. We vary the values of λ, θ and η to simulate the decision maker's different aspiration levels for the aforementioned practical investment requirements. For example, if the decision maker pays more attention to budget constraint, then we set a higher credibility level λ. We apply the solution approach to solve the proposed two models under these cases. Sensitivity comparisons between the two models with preference weight $\omega = 0.5$ are conducted to demonstrate their difference. Table 5 displays the corresponding comparative computational results about the selected strategic goals, project size, total borrowed capital and terminal wealth.

Table 5. Optimal investment strategies obtained by the two proposed models under different credibility levels.

(P_i)	(λ,θ,η)	Projects selected			Project size	f_{TBC}	$E(W_T)$
		Strategy 1	Strategy 2	Strategy 3			
	(0.60,0.60,0.60)	{P18, P19, P20}	{P5, P10, P15}	{P2, P3, P12}	9	8983.216672	15157.2716
	(0.70,0.70,0.70)	{P17, P18, P19, P20}	{P5, P10}	{P4, P11, P12}	9	9150.812764	15125.4291
	(0.80,0.80,0.80)	{P17, P18, P19, P20}	{P5, P10}	{P4, P11, P12}	9	10529.044898	16158.6654
(P_1)	(0.90,0.90,0.90)	{P17, P18, P19,P20}	{P5, P10}	{P4, P11, P12}	9	10354.080077	14819.4638
	(0.80,0.85,0.90)	{P18, P19, P20}	{P5, P15, P16}	{P3, P4, P12}	9	9958.543355	14960.2602
	(0.80,0.95,0.85)	{P14, P18, P19}	{P5, P6, P15}	{P4, P11, P12}	9	8481.363105	12251.4341
	(0.80,0.90,0.70)	{P18, P19,P20}	{P5, P10, P15}	{P4, P7, P12}	9	9407.33154	15017.2211
	(0.60,0.60,0.60)	{P18, P19, P20}	{P5, P8, P10}	{P2, P4, P7, P11}	10	8775.562092	13706.3433
	(0.70,0.70,0.70)	{P14, P18, P19}	{P5, P10, P15}	{P2, P3, P7, P11}	10	9128.185361	13593.3182
	(0.80,0.80,0.80)	{P17, P18, P19, P20}	{P5, P8}	{P2, P4, P11, P12}	10	9449.218183	13518.6544
(P_2)	(0.90,0.90,0.90)	{P13, P19}	{P5, P8, P10}	{P1, P3, P4, P11, P12}	10	7906.266587	7960.2583
	(0.80,0.85,0.90)	{P17, P19, P20}	{P5, P8, P10}	{P1, P2, P3, P12}	10	9148.113091	12933.335
	(0.80,0.95,0.85)	{P17}	{P5, P9, P10, P15}	{P1, P3, P4, P7, P12}	10	7129.709848	8597.7149
	(0.80,0.90,0.70)	{P14, P19, P20}	{P10, P15, P16}	{P2, P3, P7, P12}	10	10381.14461	14656.8229

From Table 5, we can find that the decision maker's aspiration levels for the practical investment requirements have a significant impact on R&D portfolio selection. The larger the values of λ, θ and η become, the more prudent the decision maker is. For any given (λ,θ,η), the investment strategies obtained by the two models are distinctly different. It indicates that cardinality constraints do affect the optimal portfolio

compositions and their performance. It can be seen from the last column of Table 5 that, under each preference case, the terminal wealth obtained by the model (P_1) is larger than the one of the model (P_2). Thus, we can conclude that the model (P_1) performs better than the model (P_2). It is a strong evidence that it is wise for the decision maker to form the best strategy by selecting not more than K R&D projects from the project pool. From Table 5, we can observe that, by using the model (P_1) to make a portfolio decision, P18 and P19 should be selected for strategy 1, P5 should be selected in strategy 2 and P 12 should be selected for strategy 3. From the above results, we find that we can obtain the different investment strategies by solving the two proposed models in which the different values of the parameters λ, θ and η are given. Through determining the values of the parameters λ, θ and η according to the investor's frame of mind, the investor may achieve a favorite R&D investment strategy. In a word, the obtained investment strategies under different credibility levels can freely reflect decision makers' different investment intention, which is the advantage of using the proposed models.

In the following, we take the model (P_1) with $\lambda = \theta = \eta = 0.8$ and $\omega = 0.5$ as an example to analyze the stability of the solution algorithm under different parameter values for thirty independent runs. The corresponding computational results are displayed in Table 6. To compare these results, we calculate the average relative error (ARE) as follows

$$\text{ARE}(\%) = \frac{\text{the best found average objective value} - \text{the actual found average objective value}}{\text{the best found average objective value}} \times 100\%.$$

Here, the best found average objective value is the maximum average value of all the average values obtained under different parameter cases, and the actual found average objective value is the corresponding average objective value for a certain experiment.

Table 6. Stability test results about the solution algorithm.

CS	MCN	MR	CPU times (in seconds)	Objective value	ARE(%)
40	2000	0.80	158.43	-2413.06	2.5353
40	2400	0.80	178.22	-2442.96	1.3276
40	2200	0.80	145.55	-2407.00	2.7801
40	1800	0.80	138.14	-2422.30	2.1621
40	2000	0.84	161.22	-2475.83	0.0000
40	2000	0.82	159.61	-2456.42	0.7840
40	2000	0.78	149.90	-2407.90	2.7437
40	2000	0.76	147.20	-2432.55	1.7481

From Table 6, we can find that the maximum ARE is 2.7801%, which indicates that the solution algorithm is robust with respect to the change of parameters.

6 Conclusions

In this paper, we present two fuzzy project portfolio optimization models for pharmaceutical R&D investment decision with the objectives of maximizing terminal wealth and minimizing the total borrowed capital. To the best of our knowledge, this paper is the first time to investigate R&D project portfolio selection with minimum borrowed capital in fuzzy environment. To make the proposed models more realistic, we consider several critical practical factors including borrowing constraint, budget constraint, personnel constraint, strategy selection constraint and cardinality constraint.

Considering our models with fuzzy coefficients in both the objective functions and constraints, we employ credibility theory to transform them into credibility-based chance-constrained programming problems. Then, we adopt fuzzy simulation technique to approximate the values of the aforementioned fuzzy quantities, integrate fuzzy simulation and ABC algorithm to solve our models. A numerical example is given to demonstrate the application of our models and highlight the stability of the solution algorithm. Computational results show that the proposed models are suitable for complex R&D project portfolio optimization and investors' subjective preferences can be freely incorporated into our models.

For future study, we can investigate fuzzy pharmaceutical R&D portfolio selection with several additional practical investment constraints for specific requirements such as contingency constraints, dependency constraints, flexible investment horizon, gross profit constraints, added R&D projects, budget of projects, limited budget allocation and the R&D projects from multinational companies. In addition, we will apply our models to other specific application areas and design some other effective solution algorithms to solve the complex R&D portfolio selection models.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Nos. 71971086), the Guangdong Natural Science Foundation for Distinguished Young Scholar (No. 2019B151502037), GDUPS (Liu Yongjun 2019), the Natural Science Foundation of Guangdong Province of China (No. 2018030310287) and the Fundamental Research Funds for the Central Universities (No. 2019ZD13).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Abbassi M, Ashrafi M, Tashnizi E S (2014) Selecting balanced portfolios of R&D projects with interdependencies: A cross-entropy based methodology. Technovation 34(1): 54-63.

Arratia M N M, ópez I F L, Schaeffer S E, Cruz-Reyes L (2016). Static R&D project portfolio selection in public organizations. Decis Support Syst. 84: 53-63.

Bajer D, Zorić B (2019). An effective refined artificial bee colony algorithm for numerical optimisation. Inf Sci 504: 221-275

Bansal J C, Joshi S K, Sharma H (2018). Modified global best artificial bee colony for constrained optimization problems. Comput Electr Eng. 67: 365-382.

Beaujon G J, Marin S P, McDonald G C (2001). Balancing and optimizing a portfolio of R&D projects. Nav Res Logist. 48(1): 18-40.

Bhattacharyya R, Kumar P, Kar S (2011). Fuzzy R&D portfolio selection of interdependent projects. Comput Math Appl 62(10): 3857-3870.

Biancardi M, Villani G (2017). A fuzzy approach for R&D compound option valuation. Fuzzy Sets Syst. 310: 108-121.

Bistline J E (2016). Energy technology R&D portfolio management: Modeling uncertain returns and market diffusion. Appl Energy 183: 1181-1196.

Çağlar M, Gürel S (2017). Public R&D project portfolio selection problem with cancellations. OR Spectrum 39(3): 659-687.
Carlsson C, Fullér R, Majlender P (2002). A possibilistic approach to selecting portfolios with highest utility score. Fuzzy Sets Syst. 131(1): 13-21.
Carlsson C, Fullér R, Heikkilä M, Majlender P (2007). A fuzzy approach to R&D project portfolio selection. Int J Approx Reasoning 44(2): 93-105.
Chankong V, Haimes Y Y (1983). Multiobjective Decision Making: Theory and Methodology, North-Holland, New York.
Chen L-H, Hung C-C (2010). An integrated fuzzy approach for the selection of outsourcing manufacturing partners in pharmaceutical R&D. Int J Prod Res 48(24): 7483-7506.
Chen C M, Zhu J (2011). Efficient resource allocation via efficiency bootstraps: An application to R&D project budgeting. Oper Res 59(3): 729-741.
Coffin M A, Taylor B W (1996). Multiple criteria R&D project selection and scheduling using fuzzy logic. Comput Oper Res 23(3): 207-220.
Collan M, Luukka P (2014). Evaluating R&D projects as investments by using an overall ranking from four new fuzzy similarity measure-based TOPSIS variants. IEEE Trans Fuzzy Syst 22(3): 505-515.
Copado-Méndez P J, Pozo C, Guillén-Gosálbez G, Jiménez L (2016). Enhancing the ϵ-constraint method through the use of objective reduction and random sequences: Application to environmental problems. Comput Chem Eng 87: 36-48
Deb K (2000). An efficient constraint handling method for genetic algorithms. Comput Meth Appl Mech Eng 186(2-4): 311-338.
Gemici-Ozkan B, Wu S D (2010). R&D project portfolio analysis for the semiconductor industry. Oper Res 58(6): 1548-1563.
Girotra K, Terwiesch C, Ulrich K T (2007). Valuing R&D projects in a portfolio: Evidence from the pharmaceutical industry. Manage Sci 53(9): 1452-1466.
Hassanzadeh F, Collan M, Modarres M (2012). A practical approach to R&D portfolio selection using the fuzzy pay-off method. IEEE Trans Fuzzy Syst 20(4): 615-622.
Hassanzadeh F, Modarres M, Nemati H R, Amoako-Gyampah K (2014a). A robust R&D project portfolio optimization model for pharmaceutical contract research organizations. Int J Prod Econ 158: 18-27.
Hassanzadeh F, Nemati H, Sun M (2014b). Robust optimization for interactive multiobjective programming with imprecise information applied to R&D project portfolio selection. Eur J Oper Res 238(1): 41-53.
Hesarsorkh A H, Ashayeri J, Naeini A B (2021). Pharmaceutical R&D project portfolio selection and scheduling under uncertainty: A robust possibilistic optimization approach. Comput Ind Eng 155: 107114.
Huang C C, Chu P Y, Chiang Y H (2008). A fuzzy AHP application in government-sponsored R&D project selection. Omega 36(6): 1038-1052.
Huchzermeier A, Loch C H (2001). Project management under risk: Using the real options approach to evaluate flexibility in R&D. Manage Sci 47(1): 85-101.
Karaboga D (2005). An idea based on honey bee swarm for numerical optimization. Technical Report TR06, Erciyes University, Engineering Faculty, Computer Engineering Department.
Karaboga D, Akay B (2009). A comparative study of Artificial Bee Colony algorithm. Appl Math Comput 214(1): 108-132.
Karaboga D, Akay B (2011). A modified Artificial Bee Colony (ABC) algorithm for constrained optimization problems. Appl Soft Comput 11(3): 3021-3031.
Karasakal E, Aker P (2017). A multicriteria sorting approach based on data envelopment analysis for R&D project selection problem. Omega 73: 79-92.
Marler R T, Arora J S (2010). The weighted sum method for multi-objective optimization: new insights. Struct Multidiscip Optim 41(6): 853-862.
Mikulic M (2020). Top pharmaceutical R&D projects based on net present value June 2020. https://www.statista.com/statistics/309464/major-pharmaceutical-randd-projects-based-on-net-present-value/
Mohanty R P, Agarwal R, Choudhury A K, Tiwari M K (2005). A fuzzy ANP-based approach to R&D project selection: A case study. Int J Prod Res 43(24): 5199-5216.
Nishimura J, Okada Y (2014). R&D portfolios and pharmaceutical licensing. Res Policy 43(7): 1250-1263.
Li X, Yang G (2016). Artificial bee colony algorithm with memory. Appl Soft Comput 41: 362-372.

Liu B D (2002). Theory and Practice of Uncertain Programming, Physica-Verlag, Heidelberg.
Liu B, Iwamura K (1998). Chance constrained programming with fuzzy parameters. Fuzzy Sets Syst 94(2): 227-237.
Liu B, Liu Y K (2002). Expected value of fuzzy variable and fuzzy expected value models. IEEE Trans. Fuzzy Syst 10(4): 445-450.
Pennings E, Sereno L (2011). Evaluating pharmaceutical R&D under technical and economic uncertainty. Eur J Oper Res 212(2): 374-385.
Ringuest J L, Graves S B, Case R H (2000). Conditional stochastic dominance in R&D portfolio selection. IEEE Trans Eng Manage 47(4): 478-484.
Rogers M J, Gupta A, Maranas C D (2002). Real options based analysis of optimal pharmaceutical research and development portfolios. IEEE Trans Eng Manage 41(25): 6607-6620.
Solak S, Clarke J P B, Johnson E L, Barnes E R (2010). Optimization of R&D project portfolios under endogenous uncertainty. Eur J Oper Res 207(1): 420-433.
Stummer C, Heidenberger K (2003). Interactive R&D portfolio analysis with project interdependencies and time profiles of multiple objectives. IEEE Trans Eng Manage 50(2): 175-183.
Sun Y H, Ma J, Fan Z P, Wang J (2008). A group decision support approach to evaluate experts for R&D project selection. IEEE Trans Eng Manage 55(1): 158-170.
Tian Q, Ma J, Liang J, Kwok R C W, Liu O (2005). An organizational decision support system for effective R&D project selection. Decis Support Syst 39(3): 403-413.
Tohumcu Z, Karasakal E (2010). R&D Project performance evaluation with multiple and interdependent criteria. IEEE Trans Eng Manage 57(4): 620-633.
Xiang W, Li Y, He R, Gao M, An M (2018). A novel artificial bee colony algorithm based on the cosine similarity. Comput Ind Eng 115: 54-68.
Wang J, Hwang W L (2007). A fuzzy set approach for R&D portfolio selection using a real options valuation model. Omega 35(3): 247-257.
Wang K, Mao Y, Chen J, Yu S (2018). The optimal research and development portfolio of low-carbon energy technologies: A study of China. J Clean Prod 1761: 1065-1077.
Wang K, Wang C K, Hu C (2005). Analytic hierarchy process with fuzzy scoring in evaluating multidisciplinary R&D projects in China. IEEE Trans Eng Manage 52(1): 119-129.
Zadeh L A (1965). Fuzzy sets. Inf Contr 8(3): 338-353.