



HAL
open science

CRM2DIM: A SAS macro for implementing the dual-agent Bayesian continual reassessment method

Mohamed Amine Bayar, Anastasia Ivanova, Gwénaél Le Teuff

► **To cite this version:**

Mohamed Amine Bayar, Anastasia Ivanova, Gwénaél Le Teuff. CRM2DIM: A SAS macro for implementing the dual-agent Bayesian continual reassessment method. *Computer Methods and Programs in Biomedicine*, 2019, 176, pp.211 - 223. <10.1016/j.cmpb.2019.04.025>. <hal-03486942>

HAL Id: hal-03486942

<https://hal.science/hal-03486942v1>

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

CRM2DIM: A SAS macro for implementing the Dual-Agent Bayesian Continual Reassessment Method

Mohamed Amine Bayar^{1,2}, Anastasia Ivanova³, Gwénaél Le Teuff^{1,2}

¹*Service de Biostatistique et d'Épidémiologie, Gustave Roussy, Villejuif, France*

²*INSERM U1018, CESP, Université Paris-Sud, Université Paris-Saclay, Villejuif, France*

³*Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA*

Abstract

Background and objective: The continual reassessment method (CRM) is a model-based dose-finding design for single-agent phase I oncology trials. With the advance of targeted therapies in oncology, more and more phase I trials investigate drug combinations rather than a single agent in order to find one or more maximum tolerated dose combinations. Several designs have been proposed for such dose-finding trials but only a few software packages are available to implement them. One of the designs is the two-dimensional Bayesian CRM proposed by Wang and Ivanova. We aimed at providing an easy-to-use program to implement this design.

Methods: We developed a new SAS macro, CRM2DIM, for implementing this design. This macro can be used to run a phase I dose-finding trial for two-drug combination and to perform simulations.

Results: We describe the program with its different features including the possibility to run a initial design (start-up rule), to use either a power or a logistic regression model with or without interaction term and to incorporate historical data. We illustrate its use through several simulation studies and a fictive trial.

Conclusions: The CRM2DIM macro provides a SAS implementation of the two-dimensional Bayesian CRM for dual-agent phase I oncology trials. It is an easy-to-use tool including several features. This SAS macro allows statisticians involved in early phases development a new tool for designing a dual-agent phase I oncology trials.

Keywords: phase I clinical trial, dose-finding, drug combinations, continual reassessment method, SAS MCMC procedure

1. Introduction

Continual reassessment method (CRM) developed by O'Quigley et al. [1] is a model-based dose-finding design for single-agent phase I oncology trials. The outcome of interest is the dose limiting toxicity (DLT) observed during a fixed time interval from the onset of treatment. The dose escalation usually starts at the lowest dose. The dose assigned to new patients is determined after fitting a dose-toxicity model to all available toxicity data. Some practical modifications of the CRM and a review of dose-finding methods can be found in Goodman et al. [2], Iasonos and O'Quigley [3], and Garrett-Mayer [4]. The Bayesian version of the CRM is more used than the likelihood-based version.

The CRM and its extensions have been implemented in R [5], SAS [6] and executable programs [7]. Several SAS macros implementing CRM, Bayesian model averaging CRM, fractional CRM and time-to-event CRM are available in Menon and Zink [8]. With the introduction of the specific SAS MCMC procedure in SAS/STAT 9.2, a more flexible Bayesian analysis is made possible. An illustration of this procedure for a Bayesian estimation of parameters in CRM for a single agent is available at <http://www.pharmasug.org/proceedings/2016/SP/PharmaSUG-2016-SP06.pdf>. The authors reported that the SAS

MCMC procedure yields similar results as the commonly used `dfcrm` R package (<https://cran.r-project.org/web/packages/dfcrm/index.html>).

With the recent development of targeted therapies in oncology, there are more phase I dose-finding oncology trials aiming to evaluate several agents simultaneously (established therapies with new targeted therapies) to find the maximum tolerated dose-combination (MTDC). Hence, several designs have been developed and compared [9, 10]. Riviere et al. [9] compared different model-based methods and did not show any major differences, but Hirakawa et al. [10] showed that models performance differ. In Table 1, we show some available non-commercial software and open-source programs developed in R, C/C++. The two-dimensional Bayesian CRM design proposed by Wang and Ivanova [11] was one of the first phase I dose-finding designs published for two-drug combination. It differs from the other model-based dose-finding methods for two-agent combination phase I trials in terms of the dose-toxicity model, the estimation method, and the dose-finding algorithm. To the best of our knowledge, this method is only implemented in C which is not widespread among the statistical community (<http://www.bios.unc.edu/aivanova>).

In this paper, we present a new SAS macro, CRM2DIM, for implementing the two-dimensional Bayesian CRM design

proposed by Wang and Ivanova [11]. In Section 2, we describe the two-dimensional Bayesian CRM and show how to estimate the model parameters. Section 3 describes the SAS macro and its related sub-macros. Section 4 illustrates how to conduct simulations and run a dual-agent phase I dose-finding trial using the CRM2DIM macro. In Section 6, we discuss the main features of this new macro and compare to existing programs.

2. Two-dimensional CRM method

The goal of a single-agent phase I dose-finding trial is to find the maximum tolerated dose (MTD) defined as a dose with the probability of toxicity equal to the target toxicity Γ , usually set at 0.2 or 0.3. When two agents are investigated there exist many MTDCs. In continuous dose space they make a contour of maximum tolerated combinations. In discrete dose space there is a maximum tolerated combination for each level of one of the agents. Therefore, the most ambitious goal of a dose-finding trial with dual-agent is to find all MTDCs. A less ambitious goal, adopted in many dual-agent trials, is to find a single MTDC. Wang and Ivanova [11] considered a dose-finding trial with two agents with discrete dose levels with the goal of finding all MTDCs, one for each level of the second agent. To achieve this goal, one can simply run several independent dose-finding trials, one for each level of the second agent. A more efficient approach proposed by Wang and Ivanova [11] is to run one trial exploring all combinations while summarizing data by using a parsimonious model for toxicity as a function of dual-agent.

2.1. Notation

Let the first agent have I dose levels s_i , $i = 1, \dots, I$, and the second agent have J dose levels t_j , $j = 1, \dots, J$, with $I \geq J$. Let $\psi(s_i, t_j, \theta)$ be the probability of toxicity $P\{Y = 1 | s_i, t_j, \theta\}$ at the combination (s_i, t_j) , where Y is a binary indicator of toxicity and θ is a vector of unknown parameters. Wang and Ivanova [11] proposed the following model for the probability of toxicity:

$$\log[1 - \psi(s_i, t_j, \theta)] = \alpha \log(1 - s_i) + \beta \log(1 - t_j) + \gamma \log(1 - s_i) \log(1 - t_j) \quad (1)$$

with $\theta = (\alpha, \beta, \gamma)$, these parameters are defined such that $\alpha > 0$, $\beta > 0$, $\gamma < 0$ to satisfy the assumptions of toxicity monotonicity $\psi(s_i, t_j, \theta) \leq \psi(s_{i'}, t_{j'}, \theta)$ when $i \leq i'$ and $j \leq j'$. Actually, the model is written using constants a_i , $i = 1, \dots, I$ and b_j , $j = 1, \dots, J$, with $I \geq J$:

$$\psi(s_i, t_j, \theta) = 1 - (1 - a_i)^\alpha (1 - b_j)^{\beta + \gamma \log(1 - a_i)} \quad (2)$$

The constants a_i , $i = 1, \dots, I$ and b_j , $j = 1, \dots, J$ are the re-scaled doses s_i and t_j and are specified before the trial. They are chosen to ensure a good calibration of the dose-toxicity model and are fixed from initial guesses of toxicity probability (called also working model or skeleton) for the combination (s_i, t_j) . In addition, we need to specify a prior for the parameter

vector $\theta = (\alpha, \beta, \gamma)$. One possible way to specify the prior is to consider three independent exponential priors with mean 1 for α , β and $(-\gamma)$. For a model without interaction, $\gamma = 0$.

2.2. Dose escalation

The dose escalation/de-escalation algorithm is described as follows. Assume that k patients have been treated in a trial and their outcomes observed. After posterior mean of the estimates of the toxicity probabilities are obtained, the $(k + 1)^{th}$ patient is treated at a permissible dose combination with the estimated probability of toxicity closest to the target Γ . If the last assignment was to the combination (s_i, t_j) , or (i, j) for simplicity, the permissible set of combinations is:

$$(i-1, j), (i+1, j), (i, j-1), (i, j+1), (i, j), (i+1, j-1), (i-1, j+1)$$

$$1 \leq i \leq I, 1 \leq j \leq J$$

Note that the combination $(i + 1, j + 1)$ is not included because moving from (i, j) to $(i + 1, j + 1)$ increases doses of both agents, possibly exposing patients to a higher risk of toxicity. The trial is stopped when the total number of patients is reached.

2.3. Start-up rule

An initial design, or a start-up rule, is usually used in single-agent phase I dose-finding trials before the first toxicity is observed. Such a rule is necessary with likelihood-based CRM and is optional with Bayesian CRM. In a dual-agent trial with the goal of finding one MTD for each level of the second agent, it is recommended to run the start-up until at least one toxicity is observed at each level of the second agent. In a single-agent trial the start-up usually starts at the lowest dose. In a dual-agent trial, the start-up rule initiates at the lowest level of both agents, i.e. $(1, 1)$. The dose of the first agent is escalated if no toxicity is observed. If at least one toxicity is observed, the start-up terminates at the current level of the second agent and experimentation moves to the next level of the second agent. If a toxicity was observed at level j of the second agent at dose i of the first agent, i.e. (i, j) , then the next dose combination is $(i - 2, j + 1)$. That is, the start-up is initiated at dose $i - 2$ at the next level $j + 1$. At the conclusion of the start-up, the dose-toxicity model is fitted to estimate the toxicity probabilities at each combination. The two-dimensional CRM starts at the lowest level of second agent at a dose combination with the probability of toxicity closest to Γ .

2.4. Incorporation of historical data

Historical data available from a completed single-agent dose-finding phase I trials can be included. For instance, if toxicity data on k patients treated with the first agent in a previous study is available, incorporating such historical data is performed using the power prior approach [18, 19]. The power prior is proportional to the likelihood of the historical data raised to the power of the weight $w \in [0, 1]$ multiplied by the initial prior. So,

Program	Authors	Functions	References and links
Software	MDACC	BOIN ToxFinder	https://biostatistics.mdanderson.org/softwaredownload/ https://biostatistics.mdanderson.org/softwaredownload/ Thall et al. [12]
R package	Riviere Mander Wages Braun	Dfcomb Pipe Pocrm Run.onesim	Riviere et al. [13] Mander and Sweeting [14] Wages and Varhegyi [15] http://faculty.virginia.edu/model-based_dose-finding Braun and Wang [16]
C	Wang		Wang and Ivanova [11]
C++	Yin		Yin and Yuan [17]

Table 1: List of the software and programs implementing statistical methods for designing a dual-agent phase I trial. MDACC: MD Anderson Cancer Center.

these data can be incorporated into the two-dimensional CRM by multiplying the likelihood of the current data by

$$\left\{ \prod_{l=1}^k [\psi(a_l^0, b_l^0 = 0, \theta)]^{y_l^0} [1 - \psi(a_l^0, b_l^0 = 0, \theta)]^{1-y_l^0} \right\}^w \quad (3)$$

with y_l^0 , the toxicity response of the l -th patient of the historical data. Several versions of the power prior have been developed according to whether w is considered as a fixed value (conditional), as an estimated value (empirical Bayes), or as an unknown value (full Bayesian approach). We implemented the conditional power prior. For a model without interaction, the power prior is:

$$\left\{ \prod_{l=1}^k [1 - (1 - a_i)^\alpha]^{y_l^0} [(1 - a_i)^\alpha]^{1-y_l^0} \right\}^w \quad (4)$$

The weight w , $0 \leq w \leq 1$, reflects the quantity of historical information we incorporate when analysing the current data. This raises the issue of commensurability between historical and actual data. When $w = 0$ no historical data are used in the parameters estimation, and all historical data are used with the current data for the parameters estimation when $w = 1$. Neuenschwander et al. [20] noted that considering w as an unknown value may be not appropriate and w should be specified before the trial.

3. The SAS macro CRM2DIM and its sub-macros

First, we describe the three SAS sub-macros (ESTIMATE, DEFINE_DOSE and REPORT), then, the main SAS macro CRM2DIM which calls these 3 sub-macros. Note that the value of the macro variables used in these sub-macros are transparent for the users. Figure 1 shows the functional block diagram of the SAS macro CRM2DIM with three main blocks: initialization, iterative process of the two-dimensional CRM and reporting.

3.1. SAS sub-macro ESTIMATE

The SAS sub-macro ESTIMATE allows estimating the posterior mean of toxicity probability for each combination of the

two agents. This sub-macro is exclusively based on the SAS MCMC procedure. We implemented the model of Wang and Ivanova [11] with or without interaction as shown in equation 2. We also implemented the 3 or 4-parameter standard logistic model used by Riviere et al. [13] (See Appendix F). The additional parameter in each model corresponds to the interaction term. For each model, the prior distribution of parameters is a normal or exponential distribution with different variances.

The SAS macro ESTIMATE creates a new SAS data set called *sum* containing the variables *p_post1_mean* to *p_postk_mean* corresponding to the posterior mean of toxicity probability for each combination of the two agents. This SAS data set is then merged with the raw data set. Refer to the macro variables *dist*, *nb_param*, *prior_dist*, *val_param* of the main SAS macro CRM2DIM for the parameterization of the modelling. The input parameter used by this macro is:

- `in=startup | simul | current_data`: specifies the SAS data set to be analysed. `in=startup` allows to analyse the data accumulated during the start-up if a start-up is specified and `in=simul` to analyse the simulated data. When running a trial, `in=current_data` allows to analyse the current data and for simplicity reason the information collected during the startup phase is saved in this SAS data set. This allows to avoid to create a SAS data set startup. Each SAS data set is continuously updated after each inclusion of a new cohort of patients.

3.2. SAS sub-macro DEFINE_DOSE

The SAS sub-macro DEFINE_DOSE identifies the next recommended dose combination from the posterior mean of toxicity probabilities estimated by the SAS sub-macro ESTIMATE. The way to identify the dose recommendation depends on whether it is for the start-up step (*mode = startup*) (see section 2.3), or for the two-dimensional CRM design (*mode = 2dim_design*) (See section 2.2) or for the final step of the trial (*mode = recommend_dose*). The generated output is a SAS data set containing the accumulated data prior to running this macro, and the difference between the posterior mean of toxicity probabilities and the target toxicity probability Γ . The input parameters used by this macro are:

- `in=startup | simul | current_data`: specifies the SAS data set containing the individual toxicity data already observed.

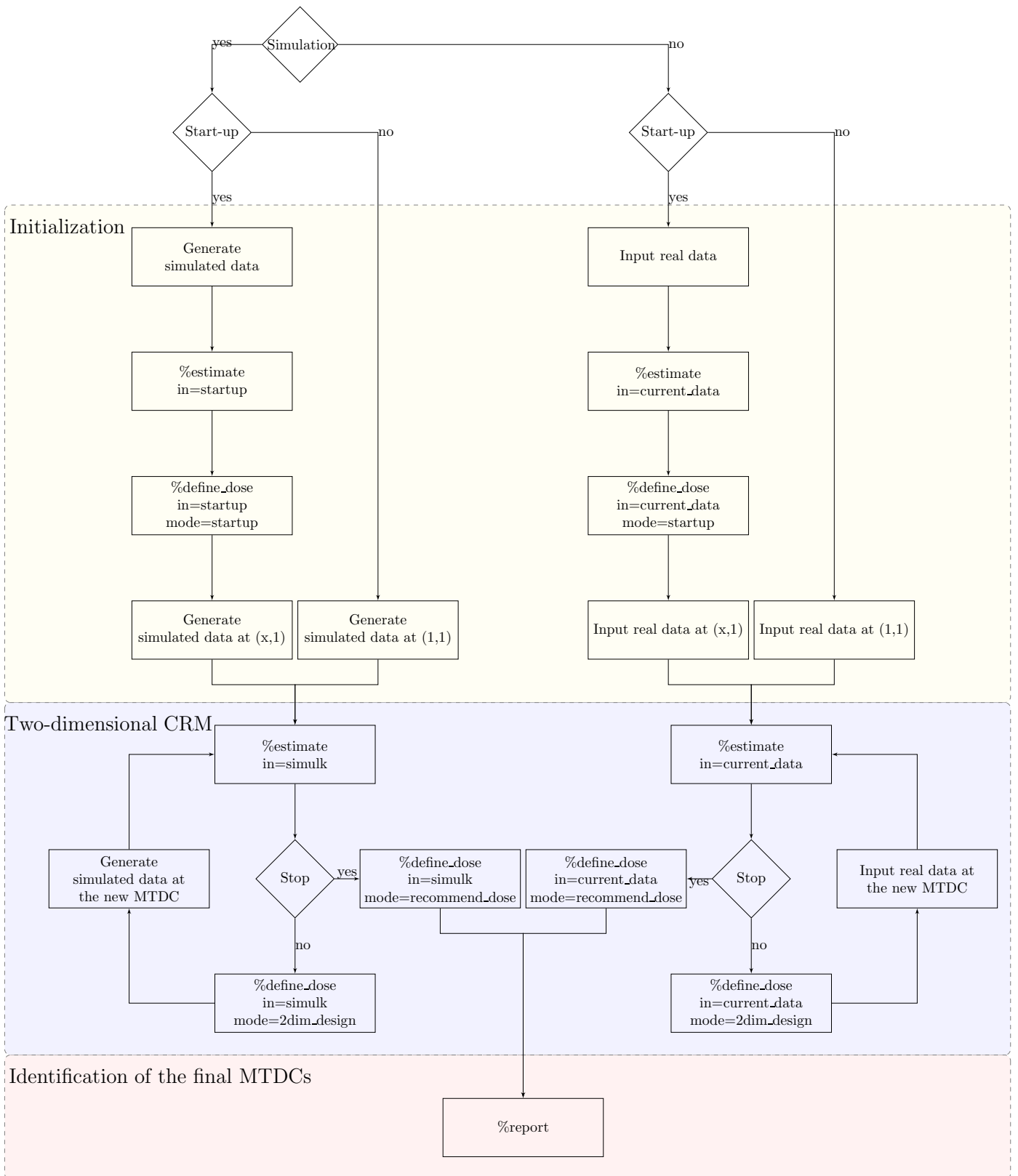


Figure 1: The functional block diagram of the SAS macro CRM2DIM. (1, 1): corresponds to dose level 1 both for the first and second agent, (x, 1): corresponds to dose level x and 1 for the first and second agent, respectively, MTDC: maximum tolerated dose-combination, simulk: simulated dataset for the k th replication

in=startup is used to analyse data accumulated during the start-up if a start-up is specified. in=simul is used to analyse simulated data. in=current_data is used to analyse current data when running a trial.

- mode=startup | 2dim_design | recommend_dose: specifies the algorithm to identify the maximum tolerated dose combination for the next cohort of patients. mode=startup indicates that the next dose is the lowest level of the second agent at a dose combination with the posterior mean of toxicity probability closest to the target toxicity probability Γ (which is defined by the macro variable gamma of the main SAS macro CRM2DIM). mode=2dim_design indicates that the recommended dose combination for the next assignment is the combination adjacent to the current one with the posterior mean of toxicity probability closest to the target toxicity probability Γ . mode=recommend_dose identifies, at the end of the trial, the MTDCs i.e drug-combination with the posterior mean of toxicity probability closest to the target toxicity probability for each level of second agent.

3.3. SAS sub-macro REPORT

The SAS sub-macro REPORT produces summary of a two-dimensional CRM for a given scenario and design parameters when simulations are run. Summary includes the percentage of experimentations at each combination, the mean number of toxicities at each combination and the percentage that a dose combination is identified as the recommended MTDC. The input parameters used by this macro are:

- in=: specifies the SAS data set containing the number of patients at each dose combination, the number of toxicities at each dose combination and the recommended dose for each dose level of the second agent for each replication of the simulation study.
- var=n_pat | n_tox | recommended_dose: specifies the variable to be analysed. var=n_pat computes the percentage of experimentations assigned to each combination. var=n_tox computes the average number of toxicities observed at each combination. var=recommended_dose computes the proportion that a dose combination is the recommended MTDC.

3.4. SAS macro CRM2DIM

The main SAS macro CRM2DIM can be used either for performing a simulation study or for conducting an actual dual-agent phase I dose-finding trial. In both cases, the program starts with the initialization of the dose level of the two agents (Figure 1 - yellow bloc) then the escalation/de-escalation starts (Figure 1 - blue bloc) and finally the identification of the final MTDCs (Figure 1 - red bloc).

When performing a simulation study (*simul* = 1), the true probabilities of toxicity at each dose combination are inputted using the macro variable *scenario*. The first cohort of patients is treated either at the lowest dose-combination (*si* = 1, *tj* = 1)

if *start_up* = 0 or at the the dose-combination resulting to the start-up rule (*start_up* = 1). In the latter case, the user has to specify the size of a cohort of simulated patients (*g_size1*) and only the dose level of the first agent (*si*) is identified since the two-dimensional CRM starts at the lowest dose level of the second agent (*tj* = 1) (See Section 2.3). The binary toxicity outcome *y* is randomly drawn from a Bernoulli distribution according to the true probability of toxicity corresponding to the initial dose-combination. For this first cohort of patients, the posterior means of toxicity probabilities at each dose combination are estimated using the SAS sub-macro ESTIMATE (see Section 3.1). The posterior means thus obtained are plugged into the SAS sub-macro DEFINE_DOSE (see Section 3.2) to define the next recommended dose-combination for a target toxicity probability level defined by Γ (macro variable *gamma*). The toxicity outcome of a new cohort of patients of size *group_size* assigned at this new recommended dose-combination is simulated and the process is repeated until the total number of patients *nb_patient* is achieved. At the end of the simulation study, a table is produced by the SAS sub-macro REPORT (See Section 3.3) and summarizes the percentage of experimentations at each combination, the mean number of toxicities at each combination and the proportion of successful dose-combination identification.

When running an actual dual-agent phase I dose-finding trial (*simul* = 0), the main difference with the simulation study is that the user has to call the SAS macro CRM2DIM each time the toxicity outcomes of the last cohort of patients treated at a certain dose-combination are collected. The macro CRM2DIM invocation utilizes the following SAS macro variables:

- simul=0 | 1: specifies if a simulation study is performed (*simul*=1) or not (*simul*=0). In the latter case, the data collected during a trial is continuously pooled by the user in a SAS data set called *current_data*
- scenario=: specifies the true toxicity probabilities at each dose combination for performing a simulation study of a two-dimensional phase I dose-finding trial. The values are entered in rows and separated by a "/" to facilitate their extraction. Each row corresponds to the true toxicity probabilities for the *I* levels of first agent at a given dose level of the second agent. This macro variable is only specified for a simulation study
- start_up=0 | 1: specified if a start-up is used (*start_up*=1) or not (*start_up*=0) (See Section 2.3). For a simulation study with a start-up, the size of each cohort is defined by the macro *g_size1*. At the end of the start-up, a recommended dose combination is identified and is assigned to the first cohort of the two-dimensional CRM. In the case of a simulation study, the size of this cohort is defined by the macro variable *group_size*
- n_patient=: specifies the total number of patients (start-up step + two-dimensional CRM) in a trial
- group_size=: specifies the number of patients per group in the two-dimensional design when a simulation study is

specified

- `nb_simul=`: specifies the number of replications in the simulation study
- `gamma=`: specifies the target toxicity probability of toxicity between 0 and 1. Usually, the targeted probability is 0.2 or 0.3
- `_a=`: specifies the set of constant $a_i (i : 1, \dots, I)$ separated by “/” for the first agent
- `_b=`: specifies the set of constant $b_j (j : 1, \dots, J)$ separated by “/” for the second agent
- `dist=0 | 1`: specifies the regression model for modelling the drug combination-toxicity relationship. The log-linear model used by Wang and Ivanova (See equations 2) is specified by `dist=0`, the standard logistic model used by [13] is specified by `dist=1`
- `nb_param=`: specifies the number of parameters for a regression model. When `dist=0` then `nb_param=2` or 3 and when `dist=1` then `nb_param=3` or 4. The core models do not include an interaction term
- `prior_dist=`: specifies the prior distribution of the model parameters. The normal prior distribution centered at 0 is symbolized by N and the exponential prior distribution by E. The prior distribution of each parameter is separated by a “/”
- `val_param=`: specifies the prior distribution variance of the model parameters. The prior distribution variance of each parameter is separated by a “/”
- `nbi=`: specifies the number of burn-in iterations for the SAS MCMC procedure
- `nmc=`: specified the number of MCMC iterations, excluding the burn-in iterations, for the SAS MCMC procedure
- `seed=`: specifies the random number seed for Markov Chain samples within the SAS MCMC procedure
- `g_size1=`: specifies the number of patients per cohort in the start-up step i.e when `start_up=1`. Usually, the values are 2 or 3 patients per cohort. `g_size1` is only specified for a simulation study
- `histo=`: specifies the SAS data set containing historical data that will be incorporated by the power prior approach (See equations 3 and 4) in the current analysis. This SAS data set requires the variables `patid`, `si` (dose level of first agent), `tj` (dose level of second agent) and `y` (a binary toxicity outcome). Its structure is one row per patient
- `a0=`: specifies the weight associated to the historical data when historical data is incorporated in the current analysis. `a0` can take the values between 0 and 1 (See Section 2.4)

	s_1	s_2	s_3	s_4	s_5	s_6
t_1	0.03	0.05	0.08	0.13	0.20	0.29
t_2	0.05	0.08	0.13	0.20	0.29	0.40
t_3	0.08	0.13	0.20	0.29	0.40	0.53

Table 2: Scenario for a two-agent phase I dose-finding trial with target toxicity probability of 0.2. Maximum tolerated combinations (or combinations closest to them) are in bold.

4. Illustration of the SAS macro CRM2DIM

4.1. Simulation

We evaluate the performance of the SAS macro CRM2DIM and its different features through a simulation study of a two drug combination phase I dose-finding trial. We used the various toxicity scenarios from Table 1 in Wang and Ivanova [11]. The first scenario of this table is described in Table 2. The rows of Table 2 represent the 3 dose levels of the second agent ($t_j, j = 1, \dots, 3$) and the columns represent the 6 dose levels of first agent ($s_i, i = 1, \dots, 6$). The target toxicity probability is 0.20 and the MTDCs are in bold. We use the same specifications as in Wang and Ivanova [11] with a total of 54 patients included in the trial ($n_{patient} = 54$) and assigning of 3 patients in each cohort ($group_size = 3$). The constants defining the working model are the same as used in O’Quigley et al. [1] for the first agent ($_a_ = 0.05/0.1/0.2/0.3/0.5/0.7$) and the second agent ($_b_ = 0.05/0.1/0.2$). In this illustration, we assume a two-parameter model (See equation 2, $dist = 0$ and $nb_param = 2$) with an exponential prior distribution ($prior_dist = E/E$) of variance equal to 1 ($val_param = 1/1$), no start-up ($start_up = 0$ and g_size1 is then left unspecified) and no inclusion of historical data ($histo$ and $a0$ are left unspecified). We ran 4000 replications ($nb_simul = 4000$) in the simulation study. The tuning of MCMC samples is $nbi = 1000$ and $nmc = 5000$ with an arbitrary $seed = 1$. The statement options $nonotes = 0$ was specified to avoid the SAS ‘log window full’ message. These input parameters are captured by the following code.

```
%CRM2DIM(simul      = 1,
          scenario    = 0.03/0.05/0.08/0.13/0.20/0.29/
                       0.05/0.08/0.13/0.20/0.29/0.40/
                       0.08/0.13/0.20/0.29/0.40/0.53,
          start_up    = 0,
          n_patient   = 54,
          group_size  = 3,
          nb_simul    = 4000,
          gamma       = 0.20,
          _a_         = 0.05/0.1/0.2/0.3/0.5/0.7,
          _b_         = 0.05/0.1/0.2,
          dist        = 0,
          nb_param    = 2,
          prior_dist  = E/E,
          val_param   = 1/1,
          nbi         = 1000,
          nmc         = 5000,
          seed        = 1,
          g_size1     = ,
          histo       = ,
          a0          =
);
```

The output generated by the SAS macro CRM2DIM with a two-parameter model, no start-up and without including a historical data is:

	Agent A Agent B	1	2	3	4	5	6
<i>pct_pat</i>	1	5.57	6.00	8.63	9.52	5.42	1.89
	2	0.32	3.23	11.10	13.25	5.11	1.48
	3	2.14	5.80	9.96	8.57	1.57	0.43
<i>n_tox</i>	1	0.08	0.15	0.35	0.65	0.56	0.27
	2	0.01	0.13	0.75	1.40	0.76	0.29
	3	0.08	0.39	1.01	1.30	0.30	0.11
<i>recommended_dose</i>	1	0.00	0.25	9.65	49.05	34.85	6.20
	2	0.03	2.23	23.88	53.35	18.20	2.33
	3	20.03	19.45	34.08	22.15	3.96	0.33

This output contains the percentages of experimentation at each combination, the average number of toxicities observed at each combination and the percentages of the MTDC recommendation. Agent A represents the first agent and agent B the second agent. The maximum tolerated dose combinations for the target toxicity probability of 0.20 for dose levels 1, 2 and 3 of second agent, the combinations (*AgentA* = 5, *AgentB* = 1), (*AgentA* = 4, *AgentB* = 2) and (*AgentA* = 3, *AgentB* = 3), are identified in 35%, 53% and 34% of simulated trials. The percentage of experimentations (average number of toxicities) at these recommended combinations are 5.4 (0.6), 13.2 (1.4) and 9.9 (1.0).

This output is produced after approximatively 7 hours using a PC Intel(R) Xeon (R) CPU E31225@ 3.10GHz, RAM: 16.0 Go, OS; Windows 7 Professional 64 bits.

To repeat this simulation study with a start-up, we set at *start_up* = 1 and specify two subjects per cohort in the start-up with *g_size1* = 2. The code becomes:

```
%CRM2DIM( simul = 1,
scenario = 0.03/0.05/0.08/0.13/0.20/0.29 /
0.05/0.08/0.13/0.20/0.29/0.40 /
0.08/0.13/0.20/0.29/0.40/0.53 ,
start_up = 1,
n_patient = 54,
group_size = 3,
nb_simul = 4000,
gamma = 0.20 ,
_a_ = 0.05/0.1/0.2/0.3/0.5/0.7 ,
_b_ = 0.05/0.1/0.2 ,
dist = 0,
nb_param = 2,
prior_dist = E/E,
val_param = 1/1,
nbi = 1000,
nmc = 5000,
seed = 1,
g_size1 = 2,
histo = ,
a0 =
);
```

The output generated by the SAS macro CRM2DIM with a two-parameter model, a start-up and no inclusion of historical data is:

	Agent A Agent B	1	2	3	4	5	6
<i>pct_pat</i>	1	3.73	4.93	7.26	10.00	6.89	2.38
	2	0.80	3.98	10.97	14.01	5.91	1.52
	3	2.12	5.59	9.34	8.15	1.88	0.54
<i>n_tox</i>	1	0.05	0.13	0.31	0.68	0.74	0.36
	2	0.02	0.16	0.75	1.46	0.90	0.32
	3	0.08	0.37	0.97	1.21	0.39	0.15
<i>recommended_dose</i>	1	0.00	0.28	10.90	49.80	32.93	6.10
	2	0.10	2.55	24.63	53.80	16.83	2.10
	3	20.03	19.95	35.70	21.00	3.10	0.23

With a start-up, the percentages of the identification of the correct maximum tolerated dose combinations, 33%, 54% and 36%, are similar compared to the design with no start-up.

We also ran simulations with a three-parameter model (See equation 2) where the third parameter is an interaction term between the two agents. This is specified by *nb_param* = 3 using the following code:

```
%CRM2DIM( simul = 1,
scenario = 0.03/0.05/0.08/0.13/0.20/0.29 /
0.05/0.08/0.13/0.20/0.29/0.40 /
0.08/0.13/0.20/0.29/0.40/0.53 ,
start_up = 1,
n_patient = 54,
group_size = 3,
nb_simul = 4000,
gamma = 0.20 ,
_a_ = 0.05/0.1/0.2/0.3/0.5/0.7 ,
_b_ = 0.05/0.1/0.2 ,
dist = 0,
nb_param = 3,
prior_dist = E/E/E,
val_param = 1/1/1,
nbi = 1000,
nmc = 5000,
seed = 1,
g_size1 = 2,
histo = ,
a0 =
);
```

The output generated by the SAS macro CRM2DIM with a three-parameter model, a start-up and no inclusion of historical data is:

	Agent A Agent B	1	2	3	4	5	6
<i>pct_pat</i>	1	3.73	4.97	7.09	11.34	8.15	2.58
	2	0.80	4.18	11.33	14.72	5.75	1.23
	3	2.25	6.21	8.74	5.37	1.16	0.39
<i>n_tox</i>	1	0.05	0.13	0.30	0.78	0.86	0.39
	2	0.02	0.18	0.79	1.54	0.88	0.27
	3	0.09	0.40	0.89	0.80	0.24	0.11
<i>recommended_dose</i>	1	0.00	0.15	9.05	42.98	40.43	7.40
	2	0.05	2.30	24.25	55.98	16.40	1.03
	3	12.48	31.50	40.70	14.38	0.95	0.00

The combinations (5,1),(4,2) and (3,3) were identified in 40%, 56% and 41% of the trials. These results indicate that the three-parameter model provide better precision of the maximum tolerated combination compared to the two-parameter model.

In addition to a simulation study with a three-parameter model and a start-up, assume that historical data is available for first agent with *k* = 10 patients treated at the combination (5,0) and

2 toxicities occurred for the last two patients. We add these k pseudo-observations made at a dose combination with level 0 of the second agent. This process is guided by the initialisation of the following variables $patid$, si , tj and y when constructing the SAS data set, $histo_data$. The following statements create this SAS data set and print the 10 historical observations.

```
data histo_data(drop=i);
  do i=1 to 8;
    patid = i;
    si = 5;
    tj = 0;
    y = 0;
    output;
  end;
  do i=1 to 2;
    patid = 8+i;
    si = 5;
    tj = 0;
    y = 1;
    output;
  end;
run;
proc print data=histo_data noobs;
run;
```

patid	si	tj	y
1	5	0	0
2	5	0	0
3	5	0	0
4	5	0	0
5	5	0	0
6	5	0	0
7	5	0	0
8	5	0	0
9	5	0	1
10	5	0	1

For a partial incorporation of the historical data, we plug-in a value of 0.5 for a_0 .

```
%CRM2DIM(simul = 1,
scenario = 0.03/0.05/0.08/0.13/0.20/0.29/
0.05/0.08/0.13/0.20/0.29/0.40/
0.08/0.13/0.20/0.29/0.40/0.53,
start_up = 1,
n_patient = 54,
group_size = 3,
nb_simul = 4000,
gamma = 0.20,
_a_ = 0.05/0.1/0.2/0.3/0.5/0.7,
_b_ = 0.05/0.1/0.2,
dist = 0,
nb_param = 3,
prior_dist = E/E/E,
val_param = 1/1/1,
nbi = 1000,
nmc = 5000,
seed = 1,
g_size1 = 2,
histo = histo_data,
a0 = 0.5
);
```

The output generated by the SAS macro CRM2DIM with a three-parameter model, a start-up and inclusion of historical data for first agent is:

	Agent A Agent B	1	2	3	4	5	6
<i>pct_pat</i>	1	3.70	3.54	6.76	14.81	8.19	1.84
	2	0.57	2.29	11.47	18.26	4.54	0.96
	3	0.91	4.38	10.29	5.99	1.10	0.38
<i>n_tox</i>	1	0.05	0.09	0.29	1.01	0.87	0.29
	2	0.01	0.09	0.79	1.91	0.69	0.21
	3	0.04	0.30	1.05	0.89	0.23	0.11
<i>recommended_dose</i>	1	0.00	0.00	1.75	51.73	43.10	3.43
	2	0.03	0.68	21.18	63.65	14.05	0.43
	3	14.38	28.88	40.48	15.55	0.73	0.00

Combination (5,1),(4,2) and (3,3) are identified in 43%, 64% and 40% after an execution time of less than 16 hours, approximatively.

We can evaluate the impact of historical data in other conditions by varying both the number of pseudo-observations k when additional historical information is available and the value of a_0 . For the former, we need to add pseudo-observations to the SAS $histo_data$ data set. For the latter, we need to change the value of a_0 according to the clinicians' prior belief about the commensurability between historical and current data.

In our simulations, we obtained almost identical results for all 5 scenarios to those reported in Wang and Ivanova [11] (See Appendix for the 4 other scenarios with a two-parameter model, a start-up and no inclusion of historical data) which validated the SAS macro, CRM2DIM.

In addition to the dose-toxicity model used by Wang and Ivanova in their two-dimensional model-based design [11], the macro includes the possibility to use a standard logistic regression for modelling the drug combination-toxicity relationship. This logistic model with 3 or 4 parameters was used in the design proposed by Riviere et al. [13] (See appendix F). The prior distributions of the model parameters were normal and exponential. The macro includes also the possibility to incorporate historical data for this model. Whatever the type of regression model ($dist = 0$ or 1 for log-linear or logistic model, respectively), val_param allows to fix different values of the variance for the different prior distributions which are specified by $prior_dist$. We illustrated these different model specifications through a simulation study with a start-up and no incorporation of historical data. We chose the scenario 1 of Table 1 in Riviere et al. [21] (with 5 and 3 dose levels for the first and second agent, respectively) and the same design parameters that those used in Riviere et al. [21] ($n_patient = 60$, $group_size = 3$, $nb_simul = 2000$, $gamma = 0.30$, $_a_ = 0.12/0.2/0.3/0.4/0.5$, $_b_ = 0.2/0.3/0.4$, $nbi = 2000$ and $nmc = 5000$). Note that the start-up rule and the dose allocation were the same as the adaptive model-based design of Wang and Ivanova [11]. The aim, here, is not to compare the 2 designs but only to illustrate the different model specifications proposed by the SAS macro CRM2DIM.

For example, the SAS code to evaluate the operating characteristics of the Wang and Ivanova design [11] but now using a 3 parameter standard logistic regression model with an intercept following a normal distribution of variance 50 and the effect of the two agents following an exponential distribution of variance 10 is as follows:

```

%CRM2DIM( simul      = 1,
          scenario    = 0.05/0.10/0.15/0.30/0.45/
                        0.10/0.15/0.30/0.45/0.55/
                        0.15/0.30/0.45/0.50/0.60,

          start_up    = 1,
          n_patient   = 60,
          group_size  = 3,
          nb_simul    = 2000,
          gamma       = 0.30,
          _a_         = 0.12/0.2/0.3/0.4/0.5,
          _b_         = 0.2/0.3/0.4,
          dist        = 1,
          nb_param    = 3,
          prior_dist  = N/E/E,
          val_param   = 50/10/10,
          nbi         = 2000,
          nmc         = 5000,
          seed        = 1,
          g_size1     = 2,
          histo       =,
          a0          =
);

```

The summary of the results for some model specifications are reported in Table 3. The columns represent the dose-toxicity model (log linear or logistic), the number of parameters (with or without interaction), the prior distribution of parameters (N and E for normal centered at 0 and exponential distribution, respectively) and some values of the variance of the prior distribution. For exponential distribution, the mean can be derived directly from the variance. For example, a variance of 1 and 10 corresponds to a mean of 1 and square root(10). The column patient represents the percentages of experimentation allocated to the true MTDC at each dose level of second agent (dose levels 1, 2, and 3), the column toxicity represents the average number of toxicities observed at the true MTDC at each dose level of second agent and the column MTDC represent the recommendation rates for true MTDC at each dose level of the second agent. In this example, the MTDC for the target toxicity probability of 0.30 for dose levels 1, 2 and 3 of second agent is (*AgentA* = 4, *AgentB* = 1), (*AgentA* = 3, *AgentB* = 2) and (*AgentA* = 2, *AgentB* = 3).

For a log-linear model (2 or 3 parameters), the results are similar for different sets of prior distributions variances. The model with an interaction (3 parameters) allows to obtain a higher recommendation rates for true MTDCs. With the logistic model, we observed some problems of convergence due to the autocorrelation of MCMC samplings and more particularly with a 4-parameter model and a high variance for prior distributions. To overcome this issue, the solutions are generally (i) to sampling by block of similar parameters i.e. putting the parameters of the model in the same block to improve the mixing of the chain, (ii) to increase the number of iterations both for burn-in period and MCMC. We thus defined a block of parameters within the SAS MCMC procedure for a logistic regression model and increase the number of burn-in and MCMC samples to reduce the autocorrelation. Using *nbi*=50000 and *nmc*=50000 allow to reduce the number of replications where autocorrelations occur for a 3-parameter logistic model and a 4-parameter logistic model with some variances fixed at 1 (less than 1%). With a 3-parameter logistic model, the recommendation rates for true MTDCs were similar to those with a 2-parameter log-

linear model but lower than those with a 3-parameter log-linear model. The 4-parameter logistic model has higher recommendations rates compared to previous models. The number of iterations both for burn-in period and MCMC would need to be higher than 50000 when evaluating the operating characteristics of 4-parameter logistic model with high variance. Indeed, a 4-parameter logistic model with variances for prior distributions of the intercept, the two dose effects and the interaction set at (50/10/10/50) lead to a high number of inconclusive replications due to autocorrelations (>20%) (data not shown).

4.2. Running a trial with the two-dimensional Bayesian CRM

As in other packages that implement dose-finding methods when running a phase I trial, the user has to alter between running a program, updating the data set after collecting the outcome of the last patients treated at a certain dose combination (*si, tj*), and running the program again until the stopping rule. To run a dual agent phase I dose-finding trial using the two-dimensional Bayesian CRM, we need to specify *simul* = 0 in the SAS macro, CRM2DIM. We remind that *g_size1* and *group_size*, corresponding to the number of patients per cohort in the start-up step and in the two-dimensional CRM, are not required since it is the user who enters the observed toxicity data for each patient. In practise, these numbers are 2 or 3 patients by cohort. The rest of the input parameters is identical to those used when performing simulation studies. If historical data are available, for example, when running a paediatric phase I trial with information from adults phase I trial then *histo_data* and *a0* need to be specified.

Suppose a fictive example designing a two-dimensional phase I dose-finding trial with 6 levels of first agent and 3 levels of second agent. The protocol specifies a target toxicity probability of 0.20 (*gamma* = 0.20), a start-up (*start_up* = 1) and a total number of patients of 35 (*n_patient* = 35). Assuming that 20 patients (10 cohorts of 2 patients) have been included in the start-up step and resulting to produce the following sequence of combinations (1,1),(2,1),(3,1),(4,1),(5,1),(6,1),(4,2),(5,2),(3,3) and (4,3) and one toxicity is observed at the combination (5,2) and another at combination (4,3). This information has to be collected in a SAS data set called *current_data*.

```

data current_data;
    input patid si tj y;
    cards;
    1 1 1 0
    2 1 1 0
    3 2 1 0
    4 2 1 0
    5 3 1 0
    6 3 1 0
    7 4 1 0
    8 4 1 0
    9 5 1 0
    10 5 1 0
    11 6 1 0
    12 6 1 0
    13 4 2 0
    14 4 2 0
    15 5 2 1
    16 5 2 0
    17 3 3 0

```

Model	Nb	Prior dist.	Variance	Patient	Toxicity	MTDC
Log-linear	2	E/E	1/1	12.23-18.50-11.48	2.11-3.26-1.98	47.00-59.10-44.45
Log-linear	2	E/E	10/10	12.01-16.72-10.89	2.11-2.91-1.88	45.45-59.10-43.90
Log-linear	3	E/E/E	1/1/1	16.27-16.71-12.85	2.87-2.91-2.28	52.95-64.10-58.25
Log-linear	3	E/E/E	10/10/10	13.67-15.44-11.43	2.38-2.70-2.03	52.30-62.00-63.75
Logistic	3	N/E/E	1/1/1	17.57-18.95-7.56	3.05-3.40-1.31	49.80-59.60-39.75
Logistic	3	N/E/E	10/10/10	19.44-20.16-9.33	3.37-3.55-1.64	49.40-59.40-42.80
Logistic	3	N/E/E	50/10/10	21.06-16.36-11.01	3.69-2.87-1.90	47.70-50.45-41.20
Logistic	4	N/E/E/N	1/1/1/1	17.96-18.37-10.91	3.14-3.26-1.93	62.30-62.80-44.30
Logistic	4	N/E/E/N	10/1/1/10	17.69-19.16-11.62	2.99-3.44-2.06	65.29-66.21-51.96

Table 3: Summary of operating characteristics in comparison of different modellings of drug combination-toxicity relationship within the two-dimensional model-based design proposed by Wang and Ivanova [11].

```

18 3 3 0
19 4 3 0
20 4 3 1
;
run ;

```

Assuming that a log-linear model with 2 parameters following an exponential prior distribution with variance set at 1 ($dist = 0$, $nb_param = 2$, $prior_dist = E/E$ and $val_param = 1/1$) is specified in the protocol. Then, the program to analyse the data produced by the start-up and thus to identify the recommended dose combination at the end of start-up is as follows:

```

%CRM2DIM(simul      = 0,
          start_up   = 1,
          n_patient  = 35,
          gamma      = 0.20,
          _a_        = 0.05/0.1/0.2/0.3/0.5/0.7,
          _b_        = 0.05/0.1/0.2,
          dist       = 0,
          nb_param   = 2,
          prior_dist = E/E,
          val_param  = 1/1,
          nbi        = 1000,
          nmc        = 5000,
          seed       = 1
        );

```

During the execution of this program, this message below is shown on the log window:

Conducting a CRM trial – start-up

The table of posterior mean of toxicity probabilities is then produced.

Agent A	1	2	3	4	5	6
Agent B						
1	0.05	0.06	0.09	0.11	0.18	0.26
2	0.09	0.10	0.12	0.15	0.21	0.29
3	0.16	0.17	0.19	0.22	0.27	0.35

Next dose combination recommendation: (5,1)

The dose combination at the first level of agent B with the estimated toxicity probability closest to the target is (5,1) ($AgentA = 5, AgentB = 1$) with the estimate of 0.18. This dose combination is the starting combination for the two-dimensional CRM design. Assuming that a cohort of three patients is thus treated at the combination (5,1) and 2 toxicities

are observed. The *current_data* data set is updated by adding three new observations with the variables $si = 5, tj = 1, y = 0$ for the first patient, $si = 5, tj = 1, y = 1$ for the second patient and $si = 5, tj = 1, y = 1$ for the third patient (without any particular order). The user needs to update the SAS data set *current_data* after inclusion of each cohort. Note that the information containing the maximum tolerated dose combination which varies during the trial is managed by two global macro variables (*current_si* and *current_tj*) and is transparent for the user. To identify the recommended dose combination for the next cohort based on data from 23 patients, 20 accrued in the start-up and 3 after the start-up, we execute the SAS macro CRM2DIM with *start_up = 0* as opposed to the analysis of data collected during the start-up. The SAS code is:

```

%CRM2DIM(simul      = 0,
          start_up   = 0,
          n_patient  = 35,
          gamma      = 0.20,
          _a_        = 0.05/0.1/0.2/0.3/0.5/0.7,
          _b_        = 0.05/0.1/0.2,
          dist       = 0,
          nb_param   = 2,
          prior_dist = E/E,
          val_param  = 1/1,
          nbi        = 1000,
          nmc        = 5000,
          seed       = 1
        );

```

The following message occurs in the log window:

Conducting a CRM trial – two dimensional design

The following output is produced i.e. the table of posterior mean of toxicity probabilities.

Agent A	1	2	3	4	5	6
Agent B						
1	0.06	0.07	0.11	0.16	0.25	0.37
2	0.09	0.11	0.15	0.19	0.28	0.39
3	0.16	0.18	0.21	0.25	0.34	0.44

Next dose combination recommendation: (4,2)

The output indicates that the next recommended combination is (4,2). This process continues until the total number of patients is reached i.e. 35 patients in this example. This means that 15 patients are thus included in the two-dimensional CRM.

Suppose we observe the following sequence of maximum tolerated dose combinations after the start-up (5,1),(4,2),(4,2),(3,3) and (4,3) with corresponding toxicity vectors (0,1,1), (0,0,1), (0,0,0), (0,0,0) and (1,1,1) for cohort of 3 patients (Figure 2). After updating the SAS data set *current_data* from the start-up ($n = 20$ patients) to the last cohort of patients in the two-dimensional CRM i.e. in pooling continuously the information which characterizes each patient i.e. (s_i, t_j, y) , we execute the same code than that executed during the two-dimensional CRM.

```
%CRM2DIM(simul = 0,
          start_up = 0,
          n_patient = 35,
          gamma = 0.20,
          _a_ = 0.05/0.1/0.2/0.3/0.5/0.7,
          _b_ = 0.05/0.1/0.2,
          dist = 0,
          nb_param = 2,
          prior_dist = E/E,
          val_param = 1/1,
          nbi = 1000,
          nmc = 5000,
          seed = 1
);
```

As the total number is reached, this execution gives automatically the recommended dose combination for each dose level of second agent. The following message occurs in the log window

Conducting a CRM trial – final recommended doses

and the following output is produced i.e. the table of posterior mean of toxicity probabilities.

Agent A	1	2	3	4	5	6
Agent B						
1	0.07	0.09	0.13	0.18	0.28	0.41
2	0.11	0.14	0.18	0.22	0.32	0.44
3	0.20	0.22	0.26	0.30	0.39	0.50

Next dose combination recommendation:(4,1),(4,2),(1,3)

Combinations (4,1),(4,2) and (1,3) are recommended as the maximum tolerated dose combinations with the estimated probabilities of toxicity equal to 0.18, 0.22 and 0.20, respectively. Note the difference between the footnotes for the outputs of the analysis of the start-up and the two-dimensional CRM during the dose-escalation compared to that for the output of the final analysis of the two-dimensional phase I dose-finding trial.

5. Mode of availability of the SAS macro CRM2DIM

Source code is available on GitHub (<https://github.com/Oncostat/CRM2DIM>).

6. Discussion

This article presents a new SAS macro, CRM2DIM, for implementing and simulating the two-dimensional Bayesian CRM

for phase I dose-finding trials with two agents proposed by Wang and Ivanova [11]. The SAS macro CRM2DIM offers several options such as a start-up, different dose-toxicity models with or without incorporation of historical data by using the simplest form of the power prior approach. As this macro is based on the use of the SAS MCMC procedure for Bayesian analysis, this macro benefits from the flexibility of this procedure and thus could be easily extended, for example, with accommodate various prior distributions, densities, specification of the log likelihood function. We illustrated some of these extensions by implementing the 3 or 4-parameter standard logistic regression used by Riviere et al. [21] with different values of the variance for the prior distributions of the model parameters. In addition, one can easily change the algorithm for dose-assignment, the start-up rule or implement stopping rules based on the toxicity posterior probability. The SAS macro CRM2DIM shares some features with the other programs reported in Table 1 for dual-agent phase I trials. Note that the approach of Hirakawa et al. [22] using a SAS/IML program for implementing the likelihood-based dose-finding method using a shrinkage logistic model is not reported since it is not available on the Web. While the R packages pocrm, pipe, dfcomb and Braun and Wang [16] and Yin and Yuan [17] programs implement dose-finding methods that identify a single MTDC, Wang and Ivanova [11] program, CRM2DIM, BOIN, and ToxFinder identify multiple MTDCs. The R packages pocrm, dfcomb, Wang and Ivanova [11] and Yin and Yuan [17] programs and CRM2DIM allow the dose-escalation method to be used with a start-up phase. A stopping rule based on posterior probability higher than a pre-specified value is implemented in dfcomb, pipe and BOIN or a fixed maximum number of patients can be enrolled or a certain number of patients treated on any combination. Only the program of Yin and Yuan [17] does not allow to run a combination phase I trial. The R package pipe offers a shiny web interface for plotting, ToxFinder also generates plots. The main feature of the SAS macro CRM2DIM compared to other programs is the incorporation of historical data. In this current version, the SAS macro CRM2DIM generates no graphic when performing simulation. We believe that this new SAS macro CRM2DIM is the first available SAS macro implementing a phase I dose-finding design for dual-agent. This macro allows to enrich the SAS toolbox dedicated to the phase I dose-finding trial in oncology.

This SAS macro CRM2DIM enables statisticians involved in early phase development to implement this two-dimensional Bayesian CRM dose-finding design in the SAS environment commonly used in the pharmaceutical companies and in some academic centres. We hope that this work will stimulate the implementation of other dose-finding methods in SAS, for example, dose-finding designs combining toxicity and efficacy.

Acknowledgments

A. Ivanova was supported in part by NIH grant P01 CA142538.

Flow chart of two-dimensional design with startup

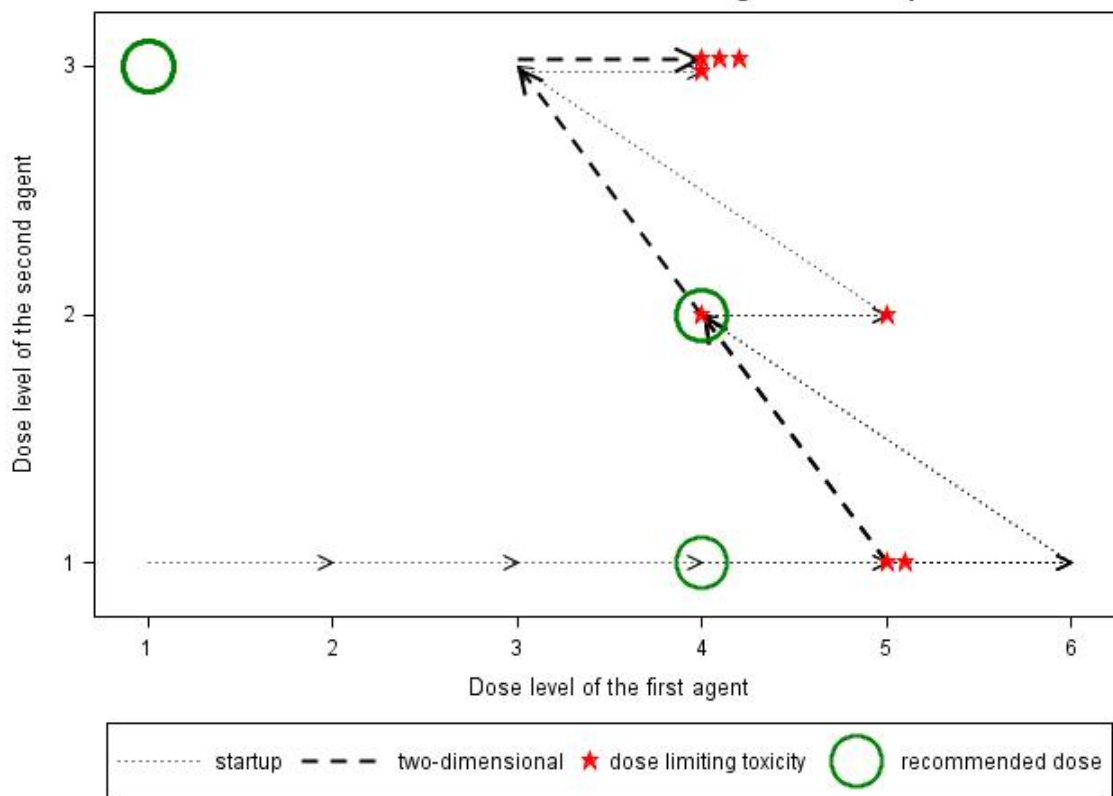


Figure 2: Example of a trial with 35 patients, 20 patients assigned in cohorts of 2 during the start-up and 15 patients in cohorts of 3 during the part guided by the two-dimensional CRM. Dotted lines represent the path observed during the start-up and dashed lines the path during the two-dimensional design. The stars correspond to observed toxicities, at combinations (5,2) and (4,3) during the start-up and (5,1),(4,2),(4,3) during the two-dimensional design. The circles indicate the maximum tolerated dose combinations recommended after the trial.

References

- [1] John O’Quigley, Margaret Pepe, and Lloyd Fisher. Continual reassessment method: A practical design for phase i clinical trials in cancer. *Biometrics*, pages 33–48, 1990.
- [2] Steven N Goodman, Marianna L Zahurak, and Steven Piantadosi. Some practical improvements in the continual reassessment method for phase i studies. *Statistics in Medicine*, 14(11):1149–1161, 1995.
- [3] Alexia Iasonos and John O’Quigley. Adaptive dose-finding studies: A review of model-guided phase i clinical trials. *Journal of Clinical Oncology*, 32(23):2505–2511, 2014.
- [4] Elizabeth Garrett-Mayer. The continual reassessment method for dose-finding studies: A tutorial. *Clinical Trials*, 3(1):57–71, 2006.
- [5] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL <https://www.R-project.org/>.
- [6] SAS Institute Inc. *SAS/STAT Software, Version 9.4*. Cary, NC, 2014. URL <http://www.sas.com/>.
- [7] Amber Salter, Charity Morgan, and Inmaculada B Aban. Implementation of a two-group likelihood time-to-event continual reassessment method using sas. *Computer Methods and Programs in Biomedicine*, 121(3):189–196, 2015.
- [8] Sandeep Menon and Richard C Zink. *Modern Approaches to Clinical Trials using SAS: Classical, Adaptive, and Bayesian Methods*. SAS Institute, 2015.
- [9] M-K Riviere, F Dubois, and S Zohar. Competing designs for drug combination in phase i dose-finding clinical trials. *Statistics in Medicine*, 34(1):1–12, 2015.
- [10] Akihiro Hirakawa, Nolan A Wages, Hiroyuki Sato, and Shigeyuki Matsui. A comparative study of adaptive dose-finding designs for phase i oncology trials of combination therapies. *Statistics in Medicine*, 34(24):3194–3213, 2015.
- [11] Kai Wang and Anastasia Ivanova. Two-dimensional dose finding in discrete dose space. *Biometrics*, 61(1):217–222, 2005.
- [12] Peter F Thall, Randall E Millikan, Peter Mueller, and Sang-Joon Lee. Dose-finding with two agents in phase i oncology trials. *Biometrics*, 59(3):487–496, 2003.
- [13] Marie-Karelle Riviere, Jacques-Henri Jourdan, and Sarah Zohar. Df-comb: An r-package for phase i/ii trials of drug combinations. *Computer Methods and Programs in Biomedicine*, 125:117–133, 2016.
- [14] Adrian P Mander and Michael J Sweeting. A product of independent beta probabilities dose escalation design for dual-agent phase i trials. *Statistics in Medicine*, 34(8):1261–1276, 2015.
- [15] Nolan A Wages and Nikole Varhegyi. Pocrm: An r-package for phase i trials of combinations of agents. *Computer Methods and Programs in Biomedicine*, 112(1):211–218, 2013.
- [16] Thomas M Braun and Shufang Wang. A hierarchical bayesian design for phase i trials of novel combinations of cancer therapeutic agents. *Biometrics*, 66(3):805–812, 2010.
- [17] Guosheng Yin and Ying Yuan. Bayesian dose finding in oncology for drug combinations by copula regression. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 58(2):211–224, 2009.
- [18] Joseph G Ibrahim, Ming-Hui Chen, Yeongjin Gwon, and Fang Chen. The power prior: Theory and applications. *Statistics in Medicine*, 34(28):3724–3749, 2015.
- [19] Anna TR Legedza and Joseph G Ibrahim. Heterogeneity in phase i clinical trials: Prior elicitation and computation using the continual reassessment method. *Statistics in Medicine*, 20(6):867–882, 2001.
- [20] Beat Neuenschwander, Michael Branson, and David J Spiegelhalter. A note on the power prior. *Statistics in Medicine*, 28(28):3562–3566, 2009.

Scenarios	s_1	s_2	s_3	s_4	s_5	s_6
Scenario 1						
t_1	0.03	0.05	0.08	0.13	0.20	0.29
t_2	0.05	0.08	0.13	0.20	0.29	0.40
t_3	0.08	0.13	0.20	0.29	0.40	0.53
Scenario 2						
t_1	0.04	0.05	0.08	0.11	0.15	0.21
t_2	0.04	0.06	0.09	0.13	0.18	0.25
t_3	0.05	0.08	0.11	0.15	0.21	0.29
Scenario 3						
t_1	0.03	0.05	0.13	0.20	0.27	0.35
t_2	0.10	0.20	0.25	0.32	0.41	0.50
t_3	0.20	0.30	0.41	0.53	0.65	0.70
Scenario 4						
t_1	0.03	0.05	0.08	0.13	0.17	0.20
t_2	0.08	0.13	0.20	0.32	0.41	0.50
t_3	0.20	0.40	0.47	0.56	0.65	0.76
Scenario 5						
t_1	0.05	0.08	0.13	0.20		
t_2	0.08	0.13	0.20	0.29		
t_3	0.13	0.20	0.29	0.40		
t_4	0.20	0.29	0.40	0.53		

- [21] Marie-Karelle Riviere, Ying Yuan, Frédéric Dubois, and Sarah Zohar. A bayesian dose-finding design for drug combination clinical trials based on the logistic model. *Pharmaceutical statistics*, 13(4):247–257, 2014.
- [22] Akihiro Hirakawa, Chikuma Hamada, and Shigeyuki Matsui. A dose-finding approach based on shrunken predictive probability for combinations of two agents in phase i trials. *Statistics in Medicine*, 32(26):4515–4525, 2013.

Appendix A. Example of toxicity scenarios of two-agent phase I dose-finding trial

The table below reports the toxicity scenarios for a two-agent phase I dose-finding trial with a target toxicity probability of 0.2 from Table 1 of Wang and Ivanova’s paper [11]. Maximum tolerated combinations (or combinations closest to them) are in bold. The columns and rows represent the dose levels of first agent s and second agent t , respectively.

Appendix B. Results of dual-agent Bayesian CRM - Scenario 2

The output reports the operating characteristics of dual-agent Bayesian CRM with no interaction between the two agents and a start-up for scenario 2. The columns and rows represent the dose levels of first agent s (represented by variable agentA) and second agent t (represented by variable agentB), respectively.

	Agent A	1	2	3	4	5	6
pct_pat	Agent B						
	1	3.73	4.80	5.03	6.12	7.47	5.15
	2	0.73	3.10	5.39	8.71	10.17	5.61
n_tox	3	1.35	3.20	5.85	10.67	7.76	5.17
	1	0.07	0.13	0.22	0.36	0.59	0.57
	2	0.01	0.10	0.24	0.57	0.95	0.73
$recommended_dose$	3	0.03	0.13	0.33	0.81	0.84	0.77
	1	0.00	0.03	2.38	18.05	39.93	39.63
	2	0.00	0.30	5.05	25.85	41.60	27.20
	3	4.75	4.90	13.73	34.70	31.08	10.85

Appendix C. Results of dual-agent Bayesian CRM - Scenario 3

The output reports the operating characteristics of dual-agent Bayesian CRM with no interaction between the two agents and a start-up for scenario 3. The columns and rows represent the dose levels of first agent s (represented by variable agentA) and second agent t (represented by variable agentB), respectively.

	Agent A	1	2	3	4	5	6
pct_pat	Agent B						
	1	3.78	8.01	14.08	8.83	3.07	1.00
	2	4.43	18.21	14.35	7.09	1.50	0.35
n_tox	3	5.20	6.44	2.51	0.98	0.17	0.02
	1	0.05	0.20	0.96	0.91	0.44	0.20
	2	0.22	1.89	1.84	1.18	0.31	0.09
$recommended_dose$	3	0.54	1.00	0.53	0.27	0.06	0.01
	1	0.00	16.88	42.83	34.43	5.45	0.43
	2	7.55	40.35	39.15	11.90	0.98	0.08
	3	76.15	18.00	5.40	0.40	0.05	0.00

Appendix D. Results of dual-agent Bayesian CRM - Scenario 4

The output reports the operating characteristics of dual-agent Bayesian CRM with no interaction between the two agents and a start-up for scenario 4. The columns and rows represent the dose levels of first agent s (represented by variable agentA) and second agent t (represented by variable agentB), respectively.

	Agent A	1	2	3	4	5	6
pct_pat	Agent B						
	1	3.75	5.82	10.80	11.14	6.10	2.12
	2	1.84	11.29	16.72	12.42	3.54	0.85
n_tox	3	4.55	5.32	2.51	1.01	0.20	0.02
	1	0.05	0.15	0.45	0.75	0.55	0.22
	2	0.07	0.74	1.72	2.10	0.75	0.22
$recommended_dose$	3	0.46	1.11	0.61	0.31	0.07	0.01
	1	0.00	5.25	23.53	51.95	15.90	3.38
	2	3.88	23.93	45.73	22.83	3.23	0.43
	3	83.83	11.58	4.35	0.25	0.00	0.00

Appendix E. Results of dual-agent Bayesian CRM - Scenario 5

The output reports the operating characteristics of dual-agent Bayesian CRM with no interaction between the two agents and a start-up for scenario 5. The columns and rows represent the dose levels of first agent s (represented by variable agentA) and second agent t (represented by variable agentB), respectively.

	Agent A	1	2	3	4
<i>pct_pat</i>	Agent B				
	1	3.83	7.14	11.26	8.84
	2	2.39	13.65	13.60	8.06
	3	5.98	9.60	5.30	1.66
	4	3.14	4.13	1.08	0.33
<i>n_tox</i>	1	0.09	0.30	0.79	0.95
	2	0.09	0.93	1.39	1.21
	3	0.42	0.96	0.78	0.35
	4	0.33	0.62	0.22	0.09
	<i>recommended_dose</i>	1	0.00	6.83	43.03
2		0.78	25.33	49.00	24.90
3		46.90	32.98	17.35	2.78
4		81.30	14.68	3.60	0.43

Appendix F. Logistic model proposed by Riviere et al.

Riviere et al. [13] used the following model:

$$\begin{aligned} \text{logit} \left[\psi \left(s_i, t_j, \theta \right) \right] = & \delta + \alpha \log \left(\frac{a_i}{1 - a_i} \right) + \beta \log \left(\frac{b_j}{1 - b_j} \right) \\ & + \gamma \log \left(\frac{a_i}{1 - a_i} \right) \log \left(\frac{b_j}{1 - b_j} \right) \end{aligned}$$

with $\theta = (\alpha, \beta, \gamma, \delta)$, these parameters are defined such that $\alpha > 0$, $\beta > 0$, $\alpha + \gamma \log \left(\frac{b_j}{1 - b_j} \right) > 0$, $\beta + \gamma \log \left(\frac{a_i}{1 - a_i} \right) > 0$ to satisfy the assumptions of toxicity monotonicity.