

**Accurate hematopoietic stem cell frequency estimates by fitting multicell Poisson models substituting to the single-hit Poisson model in limiting dilution transplantation assays.**

Thierry Bonnefoix, Mary Callanan

► **To cite this version:**

Thierry Bonnefoix, Mary Callanan. Accurate hematopoietic stem cell frequency estimates by fitting multicell Poisson models substituting to the single-hit Poisson model in limiting dilution transplantation assays.: HSC quantitation with Poisson models. *Blood*, American Society of Hematology, 2010, 116 (14), pp.2472-5. 10.1182/blood-2009-10-251546 . inserm-00520297

**HAL Id: inserm-00520297**

**<https://www.hal.inserm.fr/inserm-00520297>**

Submitted on 15 Dec 2010

**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.

**Article Title:**

Accurate hematopoietic stem cell frequency estimates by fitting multi-cell Poisson models substituting to the single-hit Poisson model in limiting dilution transplantation assays

**Running Title:** HSC quantitation with Poisson models

**Scientific category:** hematopoiesis and stem cells

**Authors:**

Thierry Bonnefoix<sup>1,2,3</sup>, Mary Callanan<sup>1,2,3</sup>

1. INSERM, U823, Oncogenic Pathways in the Haematological Malignancies, Institut Albert Bonniot, Grenoble Cedex 9, France.

2. Université Joseph Fourier-Grenoble I, Faculté de Médecine, Institut Albert Bonniot, UMR-S823, Grenoble Cedex 9, France.

3. Pôle de Recherche et Pôle de Biologie, Cellular and Molecular Haematology Unit, Plateforme Hospitalière de Génétique Moléculaire des Tumeurs, Centre Hospitalier Universitaire de Grenoble, Cedex 9, France.

**Corresponding author:**

Thierry Bonnefoix

INSERM, U823, Oncogenic Pathways in the Haematological Malignancies, Institut Albert Bonniot, Grenoble Cedex 9, France.

e-mail: [thierry.bonnefoix@ujf-grenoble.fr](mailto:thierry.bonnefoix@ujf-grenoble.fr)

Phone number: +33-(0)4-76-54-95-84

Fax number: +33-(0)4-76-54-94-25

**Word counts:**

Abstract: 150

Text: 1230

References: 323

Figure/Table: 207

## Abstract

Limiting dilution transplantation assay (LDTA) is considered as the gold standard method to assess hematopoietic stem cell (HSC) content. Traditionally, HSC frequency estimates are based on the single-hit Poisson model (SHPM) which posits that one donor HSC is sufficient to generate a progeny of detectable differentiated cells above a threshold value in hosts. However, there is no clear support for this statement, and it is conceivable that more than one donor HSC may be necessary to provide detectable reconstitution in hosts above the threshold level for detection, usually 0.5% - 1% of donor-derived cells. To address this hypothesis, we evaluated the ability of a class of multi-Cell Poisson models ( $C_{\geq 1}$ PMs) to fit to LDTAs. In seven out of the 8 reanalyzed LDTAs,  $C_{\geq 1}$ PMs plausibly compete with the traditional SHPM. Model averaging across the set of plausible models gives 1.32 to 5.88 - fold increases in HSC frequencies in comparison to the SHPM.

## Introduction

Although limiting dilution transplantation assay (LDTA) in recipient animals coupled to the single-hit Poisson model (SHPM) is considered as one of the best method for quantitating hematopoietic stem cells (HSCs),<sup>1</sup> investigators should be aware of the potential problems associated with this assay.<sup>2</sup> In particular, HSC frequency estimates traditionally rely on the SHPM<sup>1, 3</sup> which posits that one donor HSC is sufficient to generate a progeny of detectable differentiated cells above a threshold value in hosts, usually considered around 1% of donor-derived cells as estimated by flow cytometry.<sup>2</sup> It turns out that the reliability of HSC frequency estimates is critically dependent on this hypothesis. In fact, it is perfectly conceivable that the progeny of one HSC may be unable to reach this limit of detection imposed by standard flow cytometric analysis. In this context, a number of recipients having received one, or even more than one, HSC(s) may be falsely classified as negative for reconstitution. This potential situation disqualifies the use of the SHPM. To address this problem, we demonstrate that it is possible to accurately quantitate HSCs, providing that the traditional SHPM is mathematically remodeled to turn to a new class of Poisson models termed multi-Cell Poisson models ( $C_{\geq 1}$ PMs) which take into account the possibility of recipients misclassified as negative. The validity of this new modeling approach is demonstrated by reassessing 8 previously published LDTAs<sup>4, 5</sup> aimed at comparing HSC frequencies in *Hoxa-9*<sup>-/-</sup> versus wild type mice (2 LDTAs)<sup>4</sup> and in Notch ligand stimulated versus unstimulated CD34<sup>+</sup> cord blood cells (6 LDTAs).<sup>5</sup> In seven out of the 8 reanalyzed LDTAs,  $C_{\geq 1}$ PMs plausibly compete with the traditional SHPM, leading to significant changes in HSC frequency estimates in comparison to the SHPM.

## Material, models and statistical analyses.

**Source of the 8 reanalyzed LDTAs.**<sup>4, 5</sup> Details of the LDTAs numbered No.1a (Hoxa-9<sup>-/-</sup> mice), 1b (wild type mice), 2a to 2f (Notch ligand stimulated CD34<sup>+</sup> cells and controls) are described in supplemental Tables S1 and S2.

## Model assumptions.

*The SHPM.* The equation of this model is

$$p_i = \exp(-f_{c1}x_i).$$

This is the first term in the Poisson series, describing the equation of the SHPM.<sup>3, 6</sup> In the above equation,  $p_i$  is considered as the proportion of the recipient population with no detectable donor-derived blood cells (negative outcome),  $f_{c1}$  is the HSC frequency estimate and  $x_i$  is the number of cells at each cell dose group  $i$  with  $i=1,2,\dots,k$ , where  $k$  is the total number of cell doses. The major assumption of the SHPM is that only animals that receive zero HSC do not produce positive outcome, whereas animals transplanted with 1,2,...,C HSC(s) will result in positive outcome.

*The class of  $C_{\geq 1}$  PMs.* The general equation of these models is

$$p_i = \exp(-f_{c>1}x_i) + \exp(-f_{c>1}x_i) \frac{(f_{c>1}x_i)^1}{1!} + \dots + \exp(-f_{c>1}x_i) \frac{(f_{c>1}x_i)^n}{n!}$$

This equation describes a series of more sophisticated Poisson models that extend the basic SHPM by including the second term, third term, ...,  $n^{\text{th}}$  term of the Poisson series in addition to the first term of the Poisson series, where  $f_{c>1}$  is the HSC frequency estimate. The above equation can be written as

$$p_i = \sum_{n=0}^{C-1} \exp(-f_{c>1}x_i) \frac{(f_{c>1}x_i)^n}{n!}, \quad C \geq 1$$

and describes the general form of the multi-cell PMs ( $C_{\geq 1}$  PMs).  $C$  appears for each model as the minimum number of HSCs necessary to promote a detectable positive outcome. A set of 20 models is tested,  $C$  taking the value 1,2,...,20. The  $C_{\geq 1}$  PMs include the SHPM ( $C=1$ ) but leave open the possibility that more than one HSC ( $C>1$ ) is necessary to give rise to a

progeny of detectable differentiated cells in transplanted animals (positive outcome). In the case where  $C > 1$ , our hypothesis is that  $C_{\geq 1}$ PMs mathematically capture the situation in which a proportion of recipients having received HSC(s), but not in sufficient number, i.e. 1 to  $(C-1)$ , have been falsely scored as negative outcome.

### **Fitting a generalized linear model for assessing the fit of the SHPM to LDTAs.<sup>7</sup>**

In a previously published paper<sup>7</sup> we advised a statistical test aimed at estimating the fit of the SHPM to the data, and based on a generalized linear modeling approach. Briefly, testing the null hypothesis that the slope  $b$  is equal to 1 ( $b = 1$ ) explores the SHPM hypothesis, and this can be performed by a standard likelihood ratio test.<sup>8</sup> Standardized deviance residuals<sup>9</sup> were also used as SHPM checking diagnostics after fitting the SHPM<sup>3</sup> to the LDTAs.

### **Fitting the $C_{\geq 1}$ PMs to LDTAs and computation of HSC frequency estimates.**

Calculations of  $C_{\geq 1}$ PM-based HSC frequencies were obtained by a standard maximum likelihood procedure applied to binomial data.<sup>9</sup> Comparison of the  $C_{\geq 1}$ PMs models was performed with Akaike's information criterion ( $AIC$ ) and the related Akaike weights, called  $w$ , grounded on an extension of the Akaike's information criterion known as information-theoretic approach applied to model uncertainty and multimodel inference.<sup>10-15</sup> Finally,  $C_{\geq 1}$ PM-based HSC frequency estimates were averaged across the set of all plausible  $C_{\geq 1}$ PMs, along to a statistical procedure that takes into account the HSC frequency estimates and the relative Akaike weights of all plausible  $C_{\geq 1}$ PMs.<sup>10, 11</sup> See supplemental methods for further details.

## **Results and discussion.**

The SHPM is rejected for the LDTAs No. 1a, 2a, 2e ( $P < 0.05$ ) and is barely acceptable for the LDTAs No. 1b, 2d ( $P < 0.1$ ) (supplemental Table S3) casting doubt on the accuracy of the previously reported HSC frequency estimates<sup>4, 5</sup> for these 5 out of the 8 LDTAs. The standardized deviance residuals under the SHPM are presented in supplemental Figure S1A-B. Overall, it can be observed that the residuals are usually positive at low cell doses and negative at high cell doses. Given the structure of the deviance residuals,<sup>9</sup> this means that the fraction of negative mice under the SHPM tends to be underestimated at low cell doses and overestimated at high cell doses. The occurrence of such a systematic pattern favours the hypothesis that the SHPM is not correct and motivates the development of alternative Poisson models. Next,  $C_{\geq 1}$ PMs are fitted to the LDFA data with  $C$  taking the values 1 to 20. Based on

the log-likelihood values, it can be computed  $AIC$ ,  $\Delta AIC_m$  (the  $AIC$  difference between the best model and a given model in the set) and weights  $w$  that represent the probability of each model to be the expected best model (supplemental methods); a representative example of calculation is given in supplemental Table S4. Based on each Akaike weight  $w$ , it can be observed that the best model probability  $w$  obtained for a given  $C_{\geq 1}$ PM is not large relative to the weights for the other competing  $C_{\geq 1}$ PMs (supplemental Figure S2, left panels). The conclusion is that there is no model from the  $C_{\geq 1}$ PMs, including the traditional SHPM, that can be considered as the best approximating model to the data, except in LDTA No.2f exhibiting  $w$  close to 1 for the model with  $C=1$ , a strong evidence that the SHPM is the expected best model. Considering this model uncertainty for 7 out of the 8 LDTAs, HSC frequency estimates  $\tilde{f}$  were computed by a model averaging procedure across the plausible  $C_{\geq 1}$ PMs, defined as the set of models with  $\Delta AIC_m \leq 10$  (supplemental Figure S2, right panels; supplemental methods).  $C_{\geq 1}$ PM-averaged HSC frequency estimates  $\tilde{f}$  are higher than HSC frequency estimates  $f_{c1}$  based on the SHPM, with 1.32 to 5.88-fold increases (Table 1). Fitted SHPM and  $C_{\geq 1}$ PMs regression lines for the 8 LDTAs are presented in Figure 1A-H, with  $c^2/df$  ratios ( $c^2$ -dispersion statistics; supplementary methods) favouring the conclusion that  $C_{>1}$ PMs better fit to the data than the SHPM. This study strongly suggests that  $C_{\geq 1}$ PMs should be routinely used to more accurately estimate HSC frequencies in LDTAs. In line with our main finding that a single HSC may not be sufficient to generate detectable hematopoietic reconstitution in recipients, single-cell transplantations performed with various HSC-enriched populations may have chronically underestimated the total HSC frequencies.<sup>16</sup>

### **Acknowledgments**

This research is financially supported by ARC (Association pour la Recherche sur le Cancer) grant 3218.

### **Author's Contribution**

Thierry Bonnefoix performed mathematical and statistical data modeling

Mary Callanan and Thierry Bonnefoix co-wrote the paper

### **Conflict of interest disclosures**

All authors concur with the submission of this manuscript and that the material submitted for publication has not been previously reported and is not under consideration for publication elsewhere.

There is no conflicting financial interest.

### **References**

1. Szilvassy SJ, Humphries RK, Lansdorp PM, Eaves AC, Eaves CJ. Quantitative assay for totipotent reconstituting hematopoietic stem cells by a competitive repopulation strategy. *Proc Natl Acad Sci U S A*. 1990;87(22):8736-8740.
2. Purton LE, Scadden DT. Limiting Factors in Murine Hematopoietic Stem Cell Assays. *Cell Stem Cell*. 2007;1(3):263-270
3. Taswell C. Limiting dilution assays for the determination of immunocompetent cell frequencies. I. Data Analysis. *J. Immunol*. 1981;126(4):1614-1619.
4. Lawrence HJ, Christensen J, Fong S, et al. Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. *Blood*. 2005;106(12):3988-3994.
5. Delaney C, Heimfeld S, Brashem-Stein C, et al. Notch-mediated expansion of human cord blood progenitor cells capable of rapid myeloid reconstitution. *Nat. Med*. 2010;16(2):232-236.
6. Porter HE, Berry RJ. The efficient design of transplantable tumour assays. *Br. J. Cancer*. 1963;17:583-595.
7. Bonnefoix T, Bonnefoix P, Verdiel P, Sotto JJ. Fitting limiting dilution experiments with generalized linear models results in a test of the single-hit Poisson assumption *J. Immunol. Methods*. 1996;194(2):113-119.
8. Hu Y, Smyth, GK. ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods*. 2009; 347(1-2):70-78.
9. Collett D. *Modelling binary data*. Boca Raton: Chapman & Hall; 2003.
10. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York: Springer; 2002.



11. Anderson DR. Model based inference in the life sciences. Springer; 2008.
12. Claeskens G, Carroll RJ. An asymptotic theory for model selection inference in general semiparametric problems. *Biometrika*. 2007;94 (2):249-265.
13. Buena F. Consistent covariate selection and post model selection inference in semiparametric regression. *Ann. Stat.* 2004;32(3):898-927.
14. Hjort NL, Claeskens G. Frequentist model average estimators. *J. Am. Stat. Assoc.* 2003;98(464):879-899.
15. Claeskens G. Model selection and model averaging. Cambridge University Press; 2008.
16. Kent DG, Copley MR, Benz C et al. Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood*. 2009;113(25):6342-6350.

**Table 1. Results of the modeling study fitting Poisson models (SHPM and  $C_{\geq 1}$ PMs) to the 8 LDTAs with HSC frequency estimates, their standard errors and 95% confidence intervals.**

LDTA No.	Cell subset	SHPM ( $C = 1$ )			$C_{\geq 1}$ PMs				$\frac{\tilde{f}}{f_{cl}}$
		$f_{cl}^{\square}$	$SE(f_{cl})^{\dagger}$	$95\%CI(f_{cl})$	$C^{\mp}$	$\tilde{f}^{\S}$	$SE(\tilde{f})^{\P}$	$95\%CI(\tilde{f})$	
1a	Hoxa-9 <sup>-/-</sup> 12 weeks	6.11X10 <sup>-6</sup> (1/163,934)	1.65X10 <sup>-6</sup>	2.88X10 <sup>-6</sup> -9.34X10 <sup>-6</sup> (1/347,222- 1/107,066)	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	3.59X10 <sup>-5</sup> (1/27,816)	1.53X10 <sup>-5</sup>	5.98X10 <sup>-6</sup> -6.59X10 <sup>-5</sup> (1/167,179- 1/15,170)	5.88
1b	Wild type 12 weeks	4.37X10 <sup>-5</sup> (1/22,883)	1.11X10 <sup>-5</sup>	2.2X10 <sup>-5</sup> -6.55X10 <sup>-5</sup> (1/45,454- 1/15,167)	1, 2, 3, 4, 5, 6, 7, 8, 9	1.85X10 <sup>-4</sup> (1/5,416)	7.62X10 <sup>-5</sup>	3.51X10 <sup>-5</sup> -3.34X10 <sup>-4</sup> (1/28,449- 1/2,993)	4.23
2a	Delta1-IgG 3 weeks	1.25X10 <sup>-4</sup> (1/8,000)	1.94X10 <sup>-5</sup>	8.73X10 <sup>-5</sup> -1.63X10 <sup>-4</sup> (1/11,454- 1/6,135)	1, 2, 3, 4	3.32X10 <sup>-4</sup> (1/3,014)	9.05X10 <sup>-5</sup>	1.54X10 <sup>-4</sup> -5.09X10 <sup>-4</sup> (1/6,477- 1/1,964)	2.66
2b	Control-IgG 3 weeks	3.79X10 <sup>-5</sup> (1/26,385)	7.22X10 <sup>-6</sup>	2.38X10 <sup>-5</sup> -5.21X10 <sup>-5</sup> (1/42,016- 1/19,193)	1, 2, 3	7.25X10 <sup>-5</sup> (1/13,791)	3.2X10 <sup>-5</sup>	1.175X10 <sup>-5</sup> -1.33X10 <sup>-4</sup> (1/85,106-1/7,503)	1.91
2c	Noncultured 3 weeks	8.18X10 <sup>-6</sup> (1/122,249)	1.36X10 <sup>-6</sup>	5.52X10 <sup>-6</sup> -1.08X10 <sup>-5</sup> (1/181,159- 1/92,592)	1, 2	1.08X10 <sup>-5</sup> (1/92,592)	4.4X10 <sup>-6</sup>	2.17X10 <sup>-6</sup> -1.94X10 <sup>-5</sup> (1/459,559-1/51,482)	1.32
2d	Delta1-IgG 9 weeks	9.91X10 <sup>-5</sup> (1/10,090)	1.54X10 <sup>-5</sup>	6.88X10 <sup>-5</sup> -1.29X10 <sup>-4</sup> (1/14,534- 1/7,751)	1, 2, 3	1.69X10 <sup>-4</sup> (1/5,904)	7.62X10 <sup>-5</sup>	1.99X10 <sup>-5</sup> -3.18X10 <sup>-4</sup> (1/50,176-1/3,136)	1.7
2e	Control-IgG 9 weeks	5.62X10 <sup>-5</sup> (1/17,793)	1.04X10 <sup>-5</sup>	3.59X10 <sup>-5</sup> -7.65X10 <sup>-5</sup> (1/27,855- 1/13,071)	1, 2, 3	1.29X10 <sup>-4</sup> (1/7,692)	3.075X10 <sup>-5</sup>	6.97X10 <sup>-5</sup> -1.9X10 <sup>-4</sup> (1/14,343-1/5,255)	2.29
2f	Noncultured 9 weeks	1.58X10 <sup>-5</sup> (1/63,291)	2.46X10 <sup>-6</sup>	2.06X10 <sup>-5</sup> -1.1X10 <sup>-4</sup> (1/48,543- 1/9,090)	1	NA	NA	NA	

$\square$  :  $f_{cl}$  is the maximum likelihood estimate of the HSC frequency obtained on fitting the SHPM to the data;  $\dagger$  :  $SE(f_{cl})$  is the conditional standard error of  $f_{cl}$ ; based on standard normal distribution of  $f_{cl}$ ,  $95\%CI(f_{cl})$  is given by the endpoints  $f_{cl} \pm 1.96 \times SE(f_{cl})$ ;  $\mp$  : value(s) of  $C$  corresponding to the set  $P$  of plausible  $C_{\geq 1}$ PMs,  $C$  being the minimum number of HSCs necessary to promote detectable repopulation in recipients (positive outcome);  $\S$  :  $\tilde{f}$  is the  $C_{\geq 1}$ PM model-averaged HSC frequency estimate in the set  $P$  of plausible  $C_{\geq 1}$ PMs;  $\P$  :  $SE(\tilde{f})$  is the unconditional standard error of  $\tilde{f}$ ;  $95\%CI(\tilde{f})$  is

given by the endpoints  $\tilde{f} \pm 1.96 \times SE(\tilde{f})$ ; see supplemental methods for further details.

NA: not available (the SHPM is clearly the best expected model).

## Figure legend

**Figure 1. Fitted SHPM and  $C_{\geq 1}$ PM - averaged regression lines for the 8 LDTAs.** Vertical axis: expected fraction, termed  $Y_i$ , of negative mice predicted by the SHPM, or expected fraction, termed  $\tilde{Y}_i$ , of negative mice predicted by model averaging across the set  $P$  of plausible  $C_{\geq 1}$ PMs ; horizontal axis: number of injected cells  $x_i$  at each cell dose group  $i$ . (A) LDTA No.1a: *Hoxa-9*<sup>-/-</sup> bone marrow cells; (B) LDTA No. 1b: wild type bone marrow cells; (C): LDTA No. 2a: Notch ligand (delta1-IgG) stimulated CD34<sup>+</sup> cells, 3 weeks posttransplant; (D): LDTA No. 2b: CD34<sup>+</sup> cultured with control IgG, 3 weeks posttransplant; (E): LDTA No. 2c: noncultured CD34<sup>+</sup>, 3 weeks posttransplant; (F): LDTA No. 2d: Notch ligand (delta1-IgG) stimulated CD34<sup>+</sup> cells, 9 weeks posttransplant; (G): LDTA No. 2e: CD34<sup>+</sup> cultured with control IgG, 9 weeks posttransplant; (H): LDTA No. 2f: noncultured CD34<sup>+</sup>, 9 weeks posttransplant; blue line: fitted SHPM regression line; red line: fitted  $C_{\geq 1}$ PM - averaged regression line; black symbols: experimental data  $(y_i / N_i, x_i)$ , where  $y_i$  is the number of mice with negative outcome,  $N_i$  is the total number of mice and  $x_i$  is the number of injected cells, at each cell dose  $i$ . The values of  $\mathbf{c}^2 / df$  ratios (Pearson  $\mathbf{c}^2$ -dispersion statistics) reported in (A) to (H) highlight that the  $C_{\geq 1}$ PM class better fit to the data than the SHPM in 7 out the 8 LDTAs: the lower value of this ratio, the better fit of the model to the data. See supplemental methods for further details.

Figure 1 (A) to (H)

