Cytometry Part A 71A:154–162 (2007)

Statistical Considerations for Enumeration of Circulating Tumor Cells

Arjan G. J. Tibbe, M. Craig Miller, and Leon W. M. M. Terstappen*
Immunicon Corporation, Huntingdon Valley, Pennsylvania

Received 28 June 2006; Revision received 22 November 2006; Accepted 28 November 2006

Background: Circulating tumor cells (CTCs) in patients with carcinomas are extremely rare. In metastatic breast cancer, the presence of $\geq$5 CTCs in 7.5 ml of blood has been associated with short survival. As this threshold has clinical implications, it is important to recognize the limitations associated with the detection and enumeration of CTCs.
Methods: Statistical analyses were performed on data generated from a multi-center clinical trial that utilized the CellSearch™ System to isolate and enumerate CTCs in 7.5 ml blood samples. The statistical issues associated with each step of the process, from blood collection to final image analysis and CTC enumeration, were determined and implemented into a model.
Results: A model describing the statistics of the different process steps that are needed for the isolation and detection of CTCs was developed. The model uses the Poisson distribution for blood collection and empirically determined distributions for the isolation and identification of CTCs. The variability between readers was identified as one of the main sources of errors responsible for the current threshold level of five CTCs.
Conclusions: Elimination of the errors made in the identification of tumor cells isolated from 7.5 ml of blood could potentially reduce the CTC threshold for the identification of patients with a poor prognosis from the current value of five CTCs to one CTC per 7.5 ml of blood. © 2007 International Society for Analytical Cytology

Key terms: rare events; circulating tumor cells; immunomagnetic enrichment

In model systems, rare event detection is often referred to as the ability to detect a single cell of a specific composition in a background of a certain number of other cells (1–3). However, in clinical applications utilizing rare event detection, the actual issues are the ability to detect the cells of interest in a limited volume and the ability to discern these cells from other cells, particularly from cell debris and other garbage (4,5). To detect cells occurring at these low frequencies reliably, a high assay efficiency and highly standardized preparation protocol are an absolute necessity. We have developed a rare cell detection system consisting of an automated immunomagnetic sample preparation system and a semi-automated fluorescence microscope (6,7). This system was used to test the hypothesis that the presence of tumor cells in blood is associated with a poor prospect for survival. In a prospective multi center study, a threshold of five circulating tumor cells (CTCs) in 7.5 ml of whole blood was established (8). The critical role of five CTCs per 7.5 ml in separating patients with a poor prognosis from those with a much better prognosis was puzzling, and no clear biological reason could be identified. In the present study, we evaluated the influence of uncertainties in the blood collection, assay efficiency, and variation in the assignment of objects as CTCs. Although the analyses were performed with a focus on the detection of CTCs, these analyses are applicable to rare event detection in general.

MATERIALS AND METHODS

Enumeration of Circulating Tumor Cells

The rare cell detection system consists of a sample preparation and cell analysis platform that have been described elsewhere in detail (6,7). In brief, ferrofluids coated with epithelial cell-specific EpCAM antibodies are used to immunomagnetically enrich epithelial cells from 7.5 ml of blood (9). The enriched samples are stained with phycoerythrin conjugated antibodies directed against cytokeratins 8, 18, and 19, an allophycocyanin conjugated antibody to CD45 and the nuclear dye DAPI. The samples are placed in the MagNest™ Cell Presentation Device and analyzed on the CellSpotter® Analyzer, a four color semi-automated fluorescence microscope (Immunicon, Huntingdon Valley, PA). Image frames covering the entire surface of the cartridge for four different fluorescence filter cubes are captured. From the captured images, a gallery of objects meeting pre-determined criteria is automatically presented in a web-enabled browser for interpretation by a trained operator who makes the

*Correspondence to: Leon WMM Terstappen, 3401 Masons Mill Road, Huntingdon Valley, PA 19006, USA.
E-mail: lterstappen@immunicon.com
Published online 2 January 2007 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/cyto.a.20369

final selection of cells. The criteria for an object to be defined as a CTC include round to oval morphology, a visible nucleus (DAPI positive), positive staining for cytokeratin, and negative staining for CD45 (6,7). Results of cell enumeration are expressed as the number of cells per 7.5 ml of blood. The performance of the assay system is described in detail elsewhere (10).

Relation Between CTCs and Survival

Blood was drawn prior to the initiation of a new line of therapy and monthly thereafter for up to six months from 177 women with metastatic breast cancer enrolled at 20 clinical centers geographically dispersed throughout the United States (8). All patients were enrolled using IRB approved protocols and provided informed consent. At the first follow up after the initiation of therapy, CTC levels were able to be determined in 164 of the patients. Using thresholds of 1–10,000 for the CTC counts, median survivals for patients with CTC levels greater than or equal to the threshold and for patients with CTC levels less than the threshold were determined.

Poisson Distribution

The Poisson distribution applies when counting randomly distributed objects (cells) in a certain interval or volume. The probability (Pr) that a volume of 7.5 ml of blood contains x CTCs is given by:

$$\Pr(x) = \frac{\mu^x}{x!} e^{-\mu} \quad (1)$$

where μ is the average number of CTCs found in aliquots of 7.5 ml of blood.

Since it is impossible to obtain an infinite number of blood samples, n, from a single patient at a single point in time, it is only feasible to provide a best estimate for μ. This changes Eq. (1) to:

$$\Pr(x) = \frac{\bar{x}_n^x}{x!} e^{-\bar{x}_n} \quad (2)$$

where $\bar{x}_n$ is the best estimate for μ given by $\bar{x}_n = \frac{1}{n}\sum_{i=1}^{i=n} x_i$ and x_i is the number of CTCs in tube i.

Assay Efficiency for CTCs

Blood from 72 normal donors was collected into CellSave® blood collection tubes (Immunicon) and 7.5 ml was transferred to one of the conical tubes provided in the CellSearch™ Circulating Tumor Cell Kit (Veridex LLC, Warren, NJ). One µl of Control Cells (Immunicon), which consists of cultured SKBR-3 breast cancer cells labeled with the fluorescent dye DiOC16, was placed on the cap of the conical tube. The number of cells on each cap was counted using a fluorescence microscope and ranged from 5 to 30 cells (average = 17). After the number of cells was determined, the cap was placed on the tube and the tube was mixed by inversion. In case cells were present on the cap that adhered to one another, the cap was discarded and a new spike was performed since one could not be sure that the cells remain together during process which would lead to errors in determining the assay efficiency. The tubes were then processed using the CellSearch™ Circulating Tumor Cell Kit according to the manufacturer's instructions. The Control Cells were discriminated from other cells by their staining on the image captured with the fluorescent filter cube tailored for DiOC16 excitation and emission. The assay efficiency was calculated by dividing the number of SKBR-3 cells detected by the number of SKBR-3 cells added.

Variability of CTC Assignment

Two sources of variability in the assignment of an event as CTC can be distinguished

1. Variations amongst readers or intra-reader variations.
2. Other deviations that might create errors in the identification of events as CTCs.

Intra-reader variation associated with the analysis of CTCs was determined using 574 blood specimens drawn from 177 women with metastatic breast cancer and processed at six different laboratories (10). The resulting image galleries from these samples were analyzed by readers at those laboratories, who were trained at the central laboratory in the interpretation of the images and in the identification of CTCs using strict criteria. The image files from these 574 blood specimens were sent to the central laboratory where they were analyzed by a reader at the central laboratory. Intra-reader variability was assessed by comparing the two independently determined CTC counts.

Other deviations that might cause errors in the classification of events as CTCs can be of any source. Examples of these are as follows:

- Non-specific labeling
- Errors in CTC definition
- Debris
- Unknown cell populations
- Others

The origin of these errors is unknown and it is practically impossible to measure these sources of variation. The model is used to determine the size of these errors.

RESULTS

Association Between Survival and Presence of CTCs in 7.5 ml of Blood

CTC levels were systematically correlated with the median survival of 164 metastatic breast cancer patients at the first follow up after initiation of a new line of therapy. The solid squares in Figure 1 represents the median survival for patients at various CTC threshold levels. The error bars indicate the 95% confidence interval. The median survival decreases with increasing CTC numbers and reaches a plateau at a threshold value of five CTCs, approximately, indicated by the vertical line. At this cutoff the Cox Hazards ratio (dotted line), also reaches a plateau. The graph clearly demonstrates that the presence of CTCs is

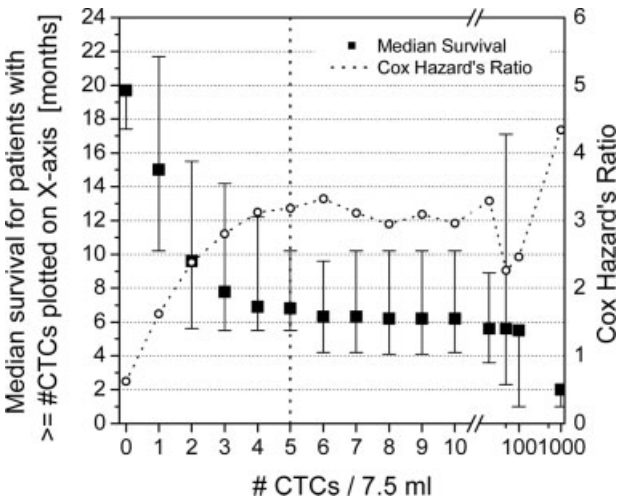

FIG. 1. The left Y-axis plus square symbols represents the median survival for metastatic breast cancer patients that had more than or equal number of tumor cells (CTCs) in a single 7.5 ml blood sample as plotted on the X-axis. The vertical line at five CTCs shows the threshold above which the survival remains at a level of 6.5 months. The right Y-axis plus dotted line shows the Cox Hazard's Ratio as a function of the number of CTCs.

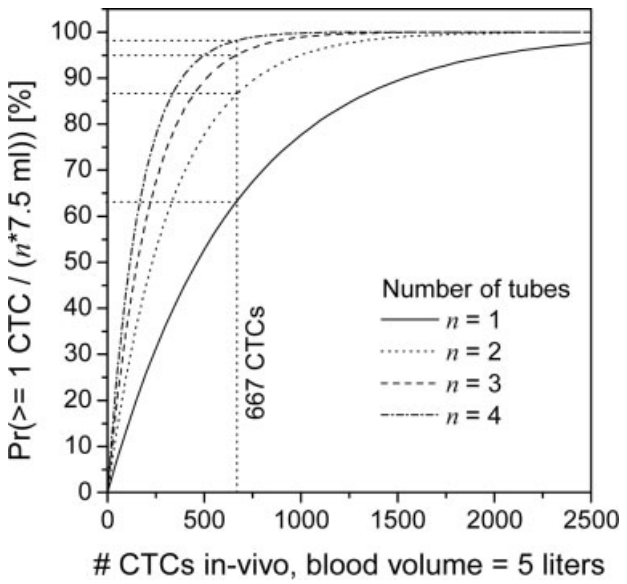

FIG. 2. The graph displays the probability of collecting ≥1 CTC in a total of n = 1, 2, 3, or 4 samples, each 7.5 ml as a function of the number of CTCs in vivo assuming a blood volume of 5 l. The solid line represents the case for n = 1.

associated with a significant decrease in survival prospects. Moreover the plateau demonstrates that patients with more than five CTCs per 7.5 ml are apparently on a futile therapy and it is therefore of great importance to know whether or not a patient falls into this category. Fifty-four of the 164 patients had five or more CTCs per 7.5 ml at their first follow-up after initiation of therapy and had a median survival of only 6.5 months.

Sampling Statistics of Blood Collection

Using formula 1 the probability of collecting ≥1 CTC in a total of n blood samples of 7.5 ml each is given by

$$\Pr(\geq 1 \text{ CTC}) = 1 - (\Pr(0))^n = 1 - e^{-n\mu} \qquad (3)$$

where μ is the average number of CTCs per 7.5 ml of blood inside the patients circulation system and n is the number of samples collected from the patient.

In Figure 2, the number of CTCs in a patient's total blood volume is plotted against the probability of collecting one or more CTCs in a total of 1, 2, 3, or 4 aliquots of 7.5 ml blood. For a patient with 667 CTCs in total blood volume of 5 l, which corresponds to one CTC per 7.5 ml, the probability of collecting ≥1 CTC in 7.5 ml of a single (n = 1) blood sample is 63%. Acquiring more samples increases the probability of at least one CTC being present in one of the n tubes and for n = 2, 3, and 4 the probability is equal to respectively 86, 95, and 98%.

Assay Efficiency for CTC Detection

The efficiency of the assay to retrieve low numbers of tumor cells from 7.5 ml blood samples was determined by spiking known numbers of cells from the breast cancer line SKBR-3 into blood samples collected from normal donors prior to processing. Figure 3 shows the results of these experiments. The number of cells spiked into each of the samples in these experiments ranged from 5 to 30, and the percent recoveries ranged from 45 to 100% (mean = 80%, standard deviation = 15%). Percent recoveries higher than 100% were not observed.

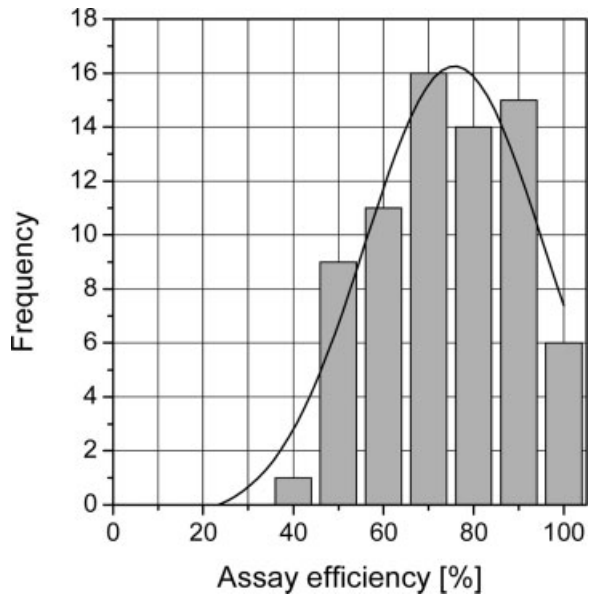

FIG. 3. Assay efficiency. Samples were processed for CTC detection after spiking 5–30 SKBR-3 cells in 7.5 ml aliquots of blood from 72 donors. The percentage of cells that were detected was determined for each experiment. The average efficiency was 80% with a standard deviation of 15%.

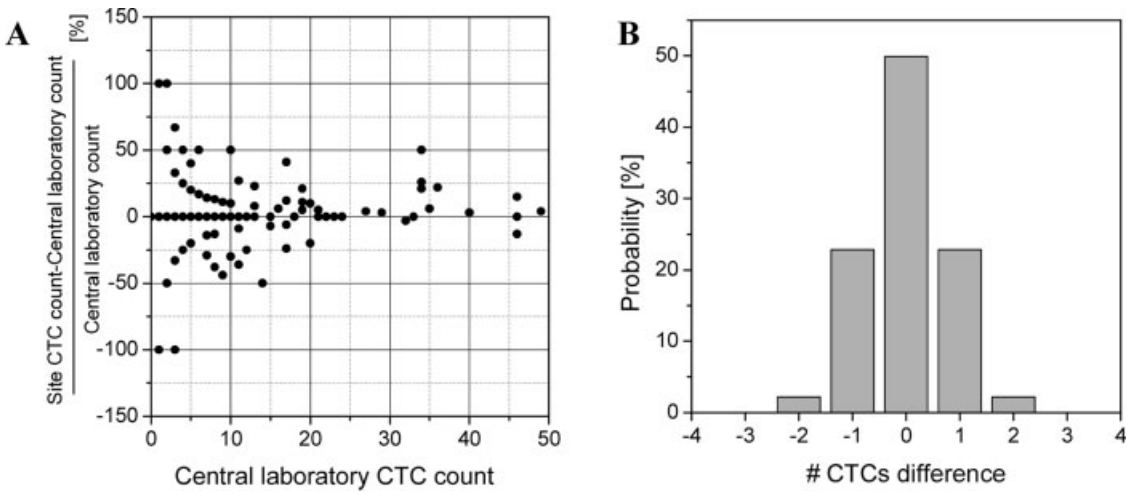

FIG. 4. **A:** Relative difference between the number of CTC detected at the site and at the Central Laboratory for samples containing <50 CTC (n = 505). **B:** The number of CTCs difference between readers that can be expected, for an intra reader standard deviation of 0.8 CTCs.

Intra-Reader Variability

The intra-reader variability was determined using the data acquired from 505 patient samples that had <50 CTCs in 7.5 ml of blood. Samples were reviewed by the site and the central laboratory (Immunicon Corporation). Every sample is reviewed twice, once by the reader of the site and once by the reader of the central laboratory. Figure 4A shows the relative difference of the number of CTCs detected by two different readers. In 277 samples the central laboratory detected zero CTCs, the Standard Deviation (SD) between site and central laboratory was 0.28. Fifty-five samples had one CTC with a SD of 0.69, 28 samples had two CTCs with a SD of 0.60, 20 samples had three CTCs with a SD of 1.20, 15 samples had four CTCs with a SD of 0.62, nine samples had five CTCs with a SD of 0.82, 8 samples had six CTCs with a SD of 1.07 and five samples had seven CTCs with a SD of 1.30. The number of samples with a CTC count larger than seven was too small to obtain an adequate standard deviation. For the range of 0–7 CTCs the standard deviation is ~0.8 CTC. For further analysis, it was assumed that this standard deviation remains equal from 0 to 15 CTCs. Figure 4B shows the difference in the number of CTC that can be expected for a standard deviation of 0.8 CTCs. In 50% of the cases, two independent readers will read the same number of tumor cells.

Statistical Model of the Analysis

In our assay 7.5 ml blood samples are drawn. The number of cells in each 7.5 ml blood sample is assumed to be Poisson distributed with an average value μ. The patient's blood volume is estimated at 5 l. A program using the graphical programming environment of Labview (National Instruments, Austin, TX) was developed to simulate the variables involved in the isolation and detection of CTCs in 7.5 ml blood samples. A flow diagram of the program that was used to simulate the statistical behavior of the process steps involved in the CTC assay is illustrated in Figure 5. The input variables are the Poisson distributed number of cells collected in 7.5 ml of blood (Fig. 2), the assay efficiency distribution (Fig. 3), and the distribution of the number of cells difference that exist between readers (Fig. 4B). The distributions of the latter two were empirically determined. The simulation starts by taking a sample out of the Poisson distribution hereby simulating the collecting of cells in the 7.5 ml blood sample during the blood draw. This number is multiplied by the sample that is taken out of the assay efficiency distribution. Since only integer numbers of CTCs exist the product of this multiplication is truncated to the next lower integer (= round to—infinity). Last, a sample is taken out of the distribution of the number of cells difference that exists between readers. This number is expressed in number of CTCs and is added to number of cells that are present after processing. The final number of detected cells is then found by truncating this number to the next lowest integer which is set to zero if negative. This sequence is repeated until the distribution of the number of cells that are detected remains unchanged (on average more than 1×10^6 times).

Simulation of the CTC Detection Process

Figure 6 shows the results of a simulation for a patient with 10,000 CTCs in its blood volume of 5 l (=15 CTCs/7.5 ml). The probability distribution for the number of CTCs collected in the tube follows the Poisson distribution, where μ = 15 CTCs/7.5 ml, is displayed as the dashed line in Figure 6. There is only a probability of 10% that exactly 15 CTCs are collected in the 7.5 ml blood sample. The simulation takes a sample out of this Poison distribution and multiplies this with a sample taken out of the probability distribution for the assay efficiency (Fig. 3). By repeating this numerous times the probability distribution of the number of CTCs that are present in the sample after processing is obtained which is presented by the dotted line in Figure 6. Addition of the intra-reader standard deviation of 0.8 CTC to the latter distribution results in the

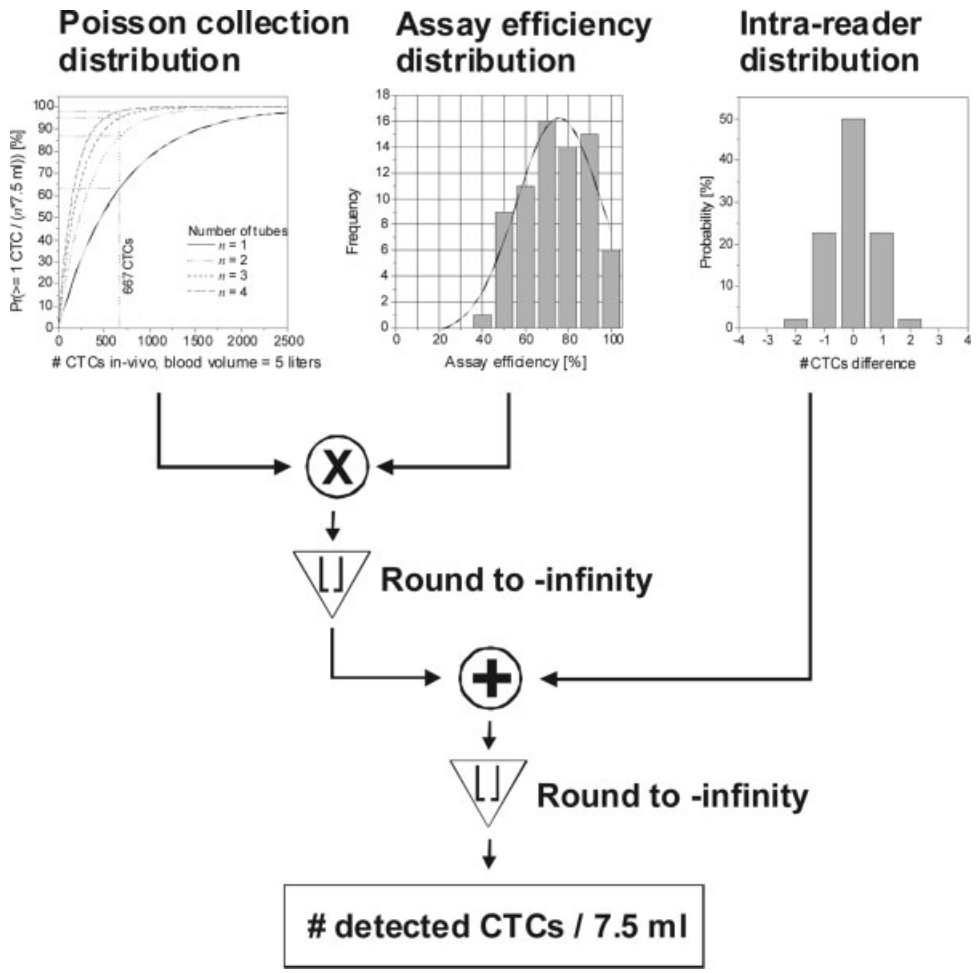

FIG. 5. Schematic representation of the program that is used to simulate the statistical behavior of the process steps involved in the assay for CTC detection.

probability distribution for the detected number of CTCs by the reader and is presented by the solid line in Figure 6. From this distribution follows that there is a probability of 4.7% that the reader detects 15 CTCs in case 10,000 CTCs are in patient's total blood volume.

The threshold level of five CTCs is indicated by the dashed vertical line. The probability that a reader will detect less than five CTCs, in case the total number of CTCs in the patient's circulation is 10,000 CTCs, is equal to

$$\Pr(<5\ \text{CTCs}) = \Pr(0\ \text{CTCs}) + \Pr(1\ \text{CTCs}) + \Pr(2\ \text{CTCs}) + \Pr(3\ \text{CTCs}) + \Pr(4\ \text{CTCs}) = 5\%$$

Although the patient in this example has many tumor cells and belongs most probably to the group of patients with a bad prognosis and short survival there is a 5% probability that this patient will be identified as a patient belonging to the group with a good prognosis. If this patient is qualified as a patient that belongs to the group with a good prognosis the median survival of this group will be reduced since this patient has a survival that is most probably shorter than the patients that have an average number of tumor cells that is lower than five. As a result the relation between survival and CTCs (Fig. 1) does contain inaccuracies and these errors will have a greater impact on those patients in which low numbers of CTCs are detected. Patients in which CTCs are identified that do not have CTCs will probably survive longer whereas patients who have no CTCs detected but do have CTCs in their circulation will possibly live shorter.

The same model is used to determine the probability of detecting $\geq$1 CTC in a single aliquot of 7.5 ml taken from a 5 l blood volume for different statistical parameters of the assay and reader. The dotted line in Figure 7A represents the theoretical situation where the assay efficiency is equal to 100% and the readers 100% correctly classified all CTCs. In this case, the probability of detecting a CTC is only limited by the Poisson statistics of the blood collection and number of samples taken. The solid line represents the probability of detecting at least one CTC using an assay efficiency of (80 $\pm$ 15)% and an intra-reader deviation of 0.8 CTCs. In this case there is an 11% probability that $\geq$1 CTC will be detected in a single 7.5 ml blood sample when there are actually no CTCs present in the 5 l of blood. This can be contributed to the false-positive identification of non-CTC-events as CTCs by the reader. The dashed line represents an assay in which the intra-reader standard deviation has been set to zero.

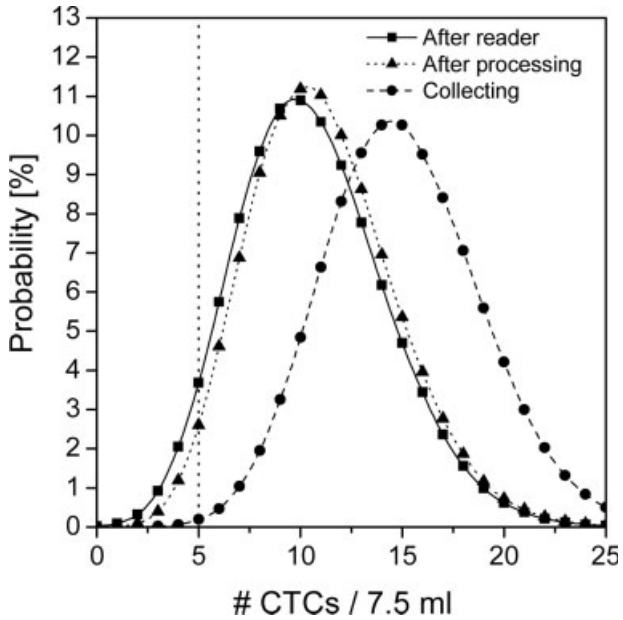

FIG. 6. Graph showing the distributions of the number of CTCs at the various steps in the process. The dashed line with circles shows the probability distribution of the number of CTCs that is collected in a single 7.5 ml blood sample from a patient that has 10,000 CTCs in its circulation assuming a total blood volume of 5 l (= 15 CTCs/7.5 ml). During processing the volume is reduced to 300 μl and the number of CTCs in this volume is distributed according to the dotted line plus triangle symbols. The reader adds an extra variation and the probability distribution of the number of CTCs that is finally detected by the reader is represented by the solid line. The threshold level of five CTCs is indicated by dotted vertical line at five CTCs per 7.5 ml.

At the level of 1,000 CTCs in 5 l of blood (or 1.5 CTCs/7.5 ml) the probability of detecting one or more CTCs in a 7.5 ml blood sample with 100% assay efficiency and no intra-reader variability (dotted line) is 77%. This is much higher compared to a probability of 43% when you have an assay efficiency of 80% without intra-reader variability (dashed line) or a probability of only 34% when you have an assay efficiency of 80% and an intra-reader variability of 0.8 CTCs (solid line).

Figure 7B displays the number of 7.5 ml samples that are needed to obtain ≥95% confidence that at least one CTC will be detected as a function of the total number of CTCs present in the patient's circulation with a total blood volume of 5 l. The solid line again represents the current assay with an efficiency of (80 ± 15)% and an intra-reader standard deviation of 0.8 CTC. The dotted line illustrates an assay efficiency of (80 ± 15)% with no intra-reader variability and the dashed line illustrates an assay efficiency of 100% with no intra-reader variability. Figure 7B shows that above 500 CTCs, the influence of the reader on the amount of samples needed is negligible. To detect ≥1 CTC with 95% confidence at the level of 1,000 CTCs in 5 l of blood (= 1.5 CTCs/7.5 ml), two aliquots are required for an assay efficiency of 100% without intra-reader variability and 6 aliquots of 7.5 ml are required in case the assay efficiency is (80 ± 15)% regardless of the intra-reader standard deviation (dotted and solid lines).

Influence of the Errors and Uncertainties on the Relation Between Number of CTCs and Survival

To determine the influence of each of the parameters in the process for the detection of CTCs one needs to start with the distribution of the number of tumor cells that circulate in the blood of the patients that were included in the study. Problem is that this is unknown. Only thing that is known is the distribution of the number of CTCs

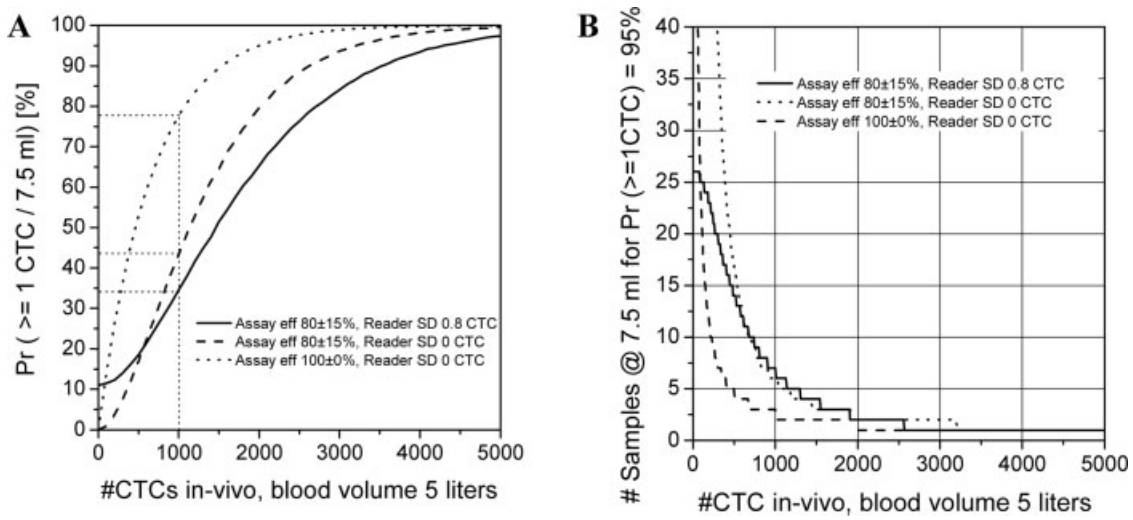

FIG. 7. **A:** The probability of detecting ≥1 CTC per 7.5 ml as a function of the number of CTCs in vivo assuming a blood volume of 5 l. The solid line represents the current assay with an assay efficiency of (80 ± 15)% and a standard deviation between readers of 0.8 CTC. The dashed and dotted line respectively display the probability for an assay efficiency of (80 ± 0.15)% with no variation between readers, assay efficiency of 100% with SD deviation between reader of zero CTCs and an assay efficiency of 100% with SD deviation between readers of 0.8 CTCs. **B:** The number of samples that need to be acquired to detect ≥1 CTC with a probability of 95%. The solid line represents again the current situation and the other lines correspond to the situations as the graph of Figure 7A.

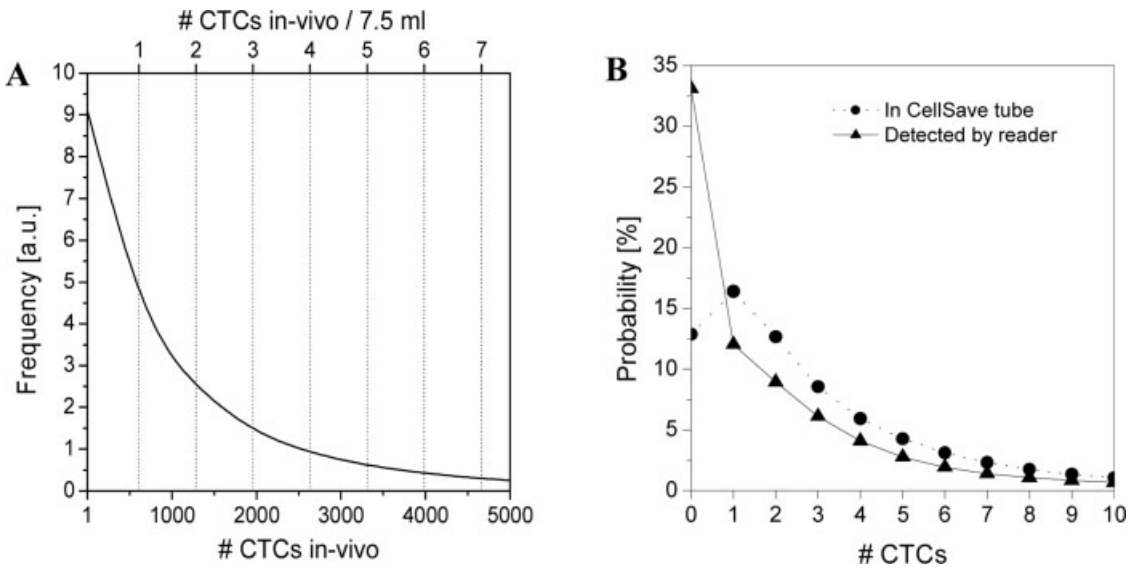

FIG. 8. **A:** The calculated in-vivo distribution of the number of CTCs per 7.5 ml for group *n*. Bottom X-axis shows the total number of CTCs in circulation; upper X-axis gives the number of CTCs *n* circulation per 7.5 ml assuming a total blood volume of 5 l. **B:** The dotted line with circles shows the probability distribution of the number of CTCs collected in the tube for group *n* in case the number of CTCs inside the bodies of the patients in this group is distributed according Figure 8A. The solid shows the probability of the number of cells that are finally detected by the reader for group *n*.

detected in the samples of these patients. The simulation was modified such that it was able to recalculate the distribution of the CTCs that resided inside the patient's blood, using the empirical determined distribution of assay and intra-reader plus Poisson statistics for the blood draw. Next to that the following assumptions were made:

- The 164 patients can be divided into a group of *m* patients without CTCs and a group of *n* patients with ≥ 1 CTCs in their blood circulation system.
- No sources for a threshold level of five CTCs could yet be identified. Therefore it is assumed that the median survival for patients in which ≥ 1 CTCs per 7.5 ml was detected, is equal to the median survival for the group in which ≥ 5 CTCs per 7.5 ml were detected. The median survival for group *n*, is 6.5 months.

The simulation minimizes the difference between the measured and simulated distribution of the number of detected CTCs and the relation between CTCs and survival by optimizing the distribution of number of CTCs that circulate in the patient's blood, the ratio of $\frac{m}{n}$ and the median survival for patients that had no CTCs in their blood, group *m*.

The minimal difference between the measured and the simulated data was obtained for a ratio of $\frac{m}{n} = 0.025$ with a distribution of the number of CTCs in circulation for patients having ≥ 1 CTCs, group *n*, as displayed in the graph of Figure 8A. The bottom X-axis presents the total number of CTCs that circulate in the blood of the patient's; the top X-axis shows the number of CTCs per 7.5 ml assuming a blood volume of 5 l. The graph drops fast, which underlines the rareness of CTCs.

A single 7.5 ml blood sample acquired from a patient of group *n*, in which the CTCs are distributed according to Figure 8A, will in 13% of the cases contain zero CTCs (dotted line Fig. 8B). After processing this 7.5 ml blood sample, with the assay efficiency and intra-reader variability's as empirically determined, there is a probability of 33% that no CTC will be detected (solid line Fig. 8B). To match the simulated median survival for the whole group of patients (*m* plus *n*) with the trial data, the median survival for group *m* needs to be set to 500 months (= 41.7 years).

The dotted line in Figure 9 displays the relation between number of CTCs and median survival based on the measured data of the 164 patients (Fig. 1). The dashed line with triangle symbols displays the median overall survival as a function of the number of CTCs detected in 7.5 ml of blood for a ratio $\frac{m}{n} = 0.025$, an assay efficiency of (80 $\pm$ 15)%, intra-reader standard deviation of 0.8 CTCs and the CTCs in the patients of group *n* distributed according to the graph of Figure 8A. The calculated median survival drops of more rapid than the measured data and reaches the plateau at ≥ 3 CTCs per 7.5 ml. As was already stated before, errors of unknown sources are not included in the calculations because their size was unknown. The dash-dotted line plus square symbols in Figure 9 shows the results in case the intra-reader standard deviation was set to 1.6 CTCs while leaving all other parameters unchanged. As is visible from the graph the calculated relation between survival and number of CTCs now exactly match. Using the simulation it is possible to determine the size of the SD of the unknown sources of variations which was in our assay for the detection of CTCs equal to 0.8 CTCs.

The solid line plus star symbol shows the relation between CTCs and survival in case the intra-reader standard deviation was set to zero. The median survival immediately drops from a median survival of 23 months at zero CTCs to the plateau level of 6.5 months at ≥ 1 CTCs.

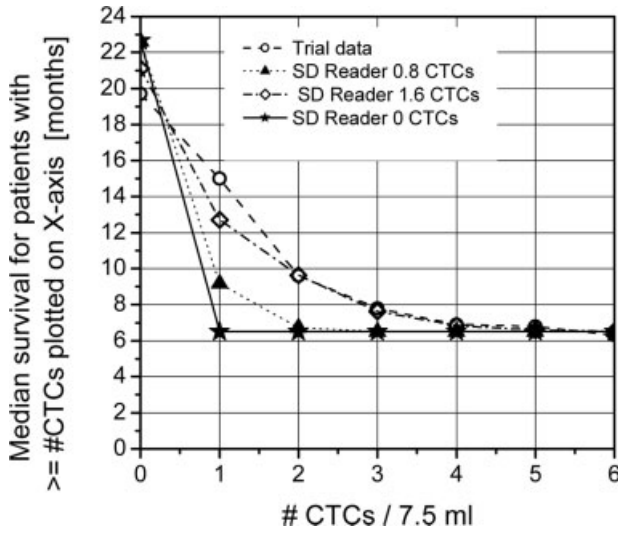

FIG. 9. The dotted line displays the relation between number of CTCs and median survival based on the measured data of the 164 patients. The dashed line with triangle symbols displays the median overall survival as a function of the number of CTCs detected in 7.5 ml of blood for a ratio $\frac{m}{n} = 0.025$, an assay efficiency of (80 ± 15)%, intra-reader standard deviation of 0.8 CTCs and the CTCs in the patients of group *n* distributed according to the graph of Figure 8A. The dash-dotted line plus square symbols presents the results in case the intra-reader standard deviation was set to 1.6 CTCs while leaving all other parameters unchanged and the solid line with star symbols shows the same situation incase the intra-reader variation was equal to zero.

From these graphs it becomes apparent that in case events can be classified as CTCs without any error, the threshold level shifts to one CTC indicating the presence of any CTC, is an identifier for a worse prognosis, independent on the exact number of CTCs.

DISCUSSION

The CellTracks technology was developed to provide a highly standardized and automated platform to detect tumor cells in whole blood. A blood volume of 7.5 ml was chosen based on reported frequencies of tumor cells in patients with metastatic disease (11–15) and the technical feasibility to immuno-magnetically select cells bearing the EpCAM antigen from this sample volume (6,7). A prospective multi-center trial conducted with this system demonstrated that the presence of circulating tumor cells in patients with metastatic breast cancer was associated with a relatively short progression free survival and survival as compared to patients in which no such cells were detected (8) (Fig. 1). Surprisingly the median overall survival did not further decrease when more than five CTCs were detected in 7.5 ml of blood. Several hypotheses can be generated to explain why this plateau was reached at five CTCs in 7.5 ml of blood. (1) CTCs need to be present at a certain concentration. (2) A subpopulation of the CTCs detected causes the decreased chances for survival, and (3) An increase in erroneous identification of CTCs when 4, 3, 2, and 1 CTCs are detected, respectively.

The graph of Figure 7B demonstrated that at a level of 1,000 in-vivo CTCs, there is a probability of 95% that at least one CTC will be detected in one out of five samples, each 7.5 ml, with the current assay. Below the level of 1,000 in-vivo CTCs the number of samples needed to detect at least one CTC with a probability of ≥95% sharply increases to a number of samples that one is unable to acquire. With the current level of detection that is achieved with CellSearch platform the limit of detection is limited by the blood volume that can be obtained from the patient.

The assumption that metastatic breast cancer patients could be separated in two distinctive groups, i.e. with and without in-vivo CTCs was used to fit the measured relation between survival and number of CTC with the calculated data. Based on this assumption the outcome of the fit demonstrated that 2.5% of the metastatic patients had truly zero in-vivo CTCs in their circulation and the median survival for this group was determined to be 500 months. In patients having CTCs, group *n*, these are distributed amongst those patients according the graph of Figure 8A.

A single 7.5 ml blood sample acquired from a patient of group *n* contains in 13% of cases no tumor cells and in 33% no CTC will be detected (Fig. 8B). From the measured median survival on the 164 metastatic breast cancer patients it is impossible to determine which patient really had no CTCs in circulation. It is therefore impossible to determine if the calculated median survival of 500 months in 2.5% of the patients is equal to the measured data. However, there are metastatic breast cancer patients with survival times of 10 years and longer.

The graph of Figure 9 showed that the measured relation between the median survival and the calculated relation exactly match if the intra-reader standard deviation was increased from 0.8 CTCs to 1.6 CTCs. This additional error 0.8 CTCs was needed to compensate for the errors of unknown sources which lead to wrong identification of events as tumor cells. In case events can be classified 100% error free, the median survival immediately drops to the plateau level of 6.5 months at one CTC. Accordingly, the threshold level of five could be replaced by a threshold level of one CTC. This suggests that the presence of a single "true" CTC in the 7.5 ml blood sample will give rise to a worse prognosis. In case one CTC per 7.5 ml is detected there are most probably many more CTCs that circulate in the blood of the patients (Fig. 2). When 500 CTCs are in circulation the probability of collecting one CTC in a 7.5 ml aliquot is 50% and the probability of detecting one CTC in this aliquot is 18%. Assuming that 500 CTCs are in circulation that are removed during one circulatory pass which takes ~60 s, a tumor should sheds 720,000 tumor cells into the blood stream during 24 h to maintain the level at 500 CTCs. Most probably one of these tumor cells will succeed to develop distant metastasis.

Addition of extra markers specific for CTCs can aid to discriminate tumor cells from other events and creates a narrower definition which will decrease the intra-reader variation and improve the identification of true CTCs. Although this will certainly help in lowering the threshold level, the limit of detection is in the end not limited

by addition of extra CTC identifiers or instrument improvement but by the amount of blood that can be examined for the presence of CTCs.

ACKNOWLEDGMENTS

The authors acknowledge, Jan Greve, David Ehus and Jonathan Uhr for their constructive criticism of the manuscript.

LITERATURE CITED

1. Gross HJ, Verwer B, Houck D, Hoffman RA, Recktenwald D. Model study detecting breast cancer cells at frequencies as low as 10-7. Proc Natl Acad Sci USA 1995;92:537–541.
2. Bajaj S, Welsh JB, Leif RC, Price JH. Ultra-rare-event detection performance of a custom scanning cytometer on a model preparation of fetal nRBCs. Cytometry 2000;39:285–294.
3. Rosenblatt JI, Hokanson JA, McLaughlin SR, Leary JF. Theoretical basis for sampling statistics useful for detecting and isolating rare cells using flow cytometry and cell sorting. Cytometry 1997;27:233–238.
4. Dumont LJ. Sampling errors and the precision associated with counting very low numbers of white cells in blood components. Transfusion 1991;31:428–431.
5. Terstappen LWMM. Detection of infrequent cells in blood and bone marrow by flowcytometry. In: Ho A, Haas R, Champlin RE, editors. Hematopoietic Stem Cell Transplantation. New York: Marcel Dekker; 2000. pp 137–152.
6. Kagan M, Howard D, Bendele T, Mayes J, Silvia J, Repollet M, Doyle J, Allard J, Tu N, Bui T, Russell T, Rao C, Hermann M, Rutner H, Terstappen LWMM. A sample preparation and analysis system for identification of circulating tumor cells. J Clin Ligand Assay 2002;25:104–110.
7. Kagan M, Howard D, Bendele T, Rao C, Terstappen LWMM. Circulating tumor cells as cancer markers, a sample preparation and analysis system. In: Diamandis EP, Fritsche HA, Lilja H, Chan DW, Schwartz M, editors. Tumor Markers: Physiology, Pathobiology, Technology and Clinical Applications. Washington: AACC Press; 2002. pp 495–498.
8. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, Rueben JM, Doyle G, Allard WJ, Terstappen LWMM, Hayes DF. Circulating tumor cells, disease progression and survival in metastatic breast cancer. N Engl J Med 2004;351:781–791.
9. Rao CG, Chianese D, Doyle GV, Miller MC, Russell T, Sanders RA, Terstappen LWMM. Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors. Int J Oncol 2005;27:49–57.
10. Allard WJ, Matera J, Miller MC, Repollet M, Connelly MC, Rao C, Tibbe A, Uhr JW, Terstappen LWMM. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with non-malignant diseases. Clin Cancer Res 2004;10:6897–6904.
11. Myerowitz RL, Edwards PA, Sartiano GP. Carcinocythemia (carcinoma cell leukemia) due to metastatic carcinoma of the breast: Report of a case. Cancer 1977;40:3107–3111.
12. Gallivan MV, Lokich JJ. Carcinocythemia (carcinoma cell leukemia). Report of two cases with English literature review. Cancer 1984;53:1100–1102.
13. Yam LT, Janckila AJ. Immunocytodiagnosis of carcinocythemia in disseminated breast cancer. Acta Cytol 1987;31:68–72.
14. Racila E, Euhus D, Weiss AJ, Rao C, McConnell J, Terstappen LWMM, Uhr JW. Detection and characterization of carcinoma cells in the blood. Proc Natl Acad Sci 1998;95:4589–4594.
15. Terstappen LWMM, Rao C, Gross S, Weiss A. Peripheral blood tumor cell load reflects the clinical activity of the disease in patients with carcinoma of the breast. Int J Oncol 2000;17:573–578.