Cancer Immunol Immunother (2013) 62:1547–1551
DOI 10.1007/s00262-013-1458-y

OPINION PAPER

Late divergence of survival curves in cancer immunotherapy trials: interpretation and implications

Fredrik B. Thorén · Harald Anderson · Örjan Strannegård

Received: 18 February 2013 / Accepted: 15 July 2013 / Published online: 24 August 2013

Abstract Late divergence of survival curves of treated patients and controls is commonly seen in successful cancer immunotherapy trials. Although late survival curve divergence may be caused by a delayed action of therapy, it may also be related to early effects of the treatment. We suggest that late survival divergence most often reflects a specific benefit of therapy for patients who suffer from a comparatively slow progression of disease. The occurrence of delayed survival curve divergence has important implications for the statistical analysis of immunotherapy trials. Thus, it leads to non-proportional hazard ratios that make commonly used statistical tests, e.g., the logrank test, suboptimal. It is therefore suggested that the statistical analysis of immunotherapy trials primarily should be based on a test that compares the survival curves at or after a prespecified, fixed, late time point.

Keywords Immunotherapy · Survival analysis · Kaplan–Meier · Treatment outcome · Late divergence · Statistical models

F. B. Thorén (✉)
Sahlgrenska Cancer Center, University of Gothenburg, Box 425, 405 30 Göteborg, Sweden
e-mail: fredrik.thoren@gu.se

F. B. Thorén · Ö. Strannegård
Department of Infectious Diseases, University of Gothenburg, Göteborg, Sweden

H. Anderson
Department of Cancer Epidemiology, University of Lund, Lund, Sweden

Introduction

In recent years, it has been recognized that survival curves of treated patients and controls obtained from successful immunotherapy trials commonly display late divergence [1–10]. Late curve divergence was initially considered to be difficult to explain [11], but the repeated observations of the phenomenon have led to its general acceptance as a distinctive and not uncommon feature of successful immunotherapy. The phenomenon has usually been interpreted as a delayed effect of the treatment, being in contrast to the immediate effect and early curve divergence seen in chemotherapy [9, 12]. For instance, treatment with the immunotherapeutic drug ipilimumab resulted in effects that were not graphically observable until several months after initiation of treatment, resulting in late divergence of the survival curves [1]. In contrast, treatment with vemurafenib, which has a mechanism of action that is seemingly independent on immunological actions, is characterized by very early, albeit transient, antitumor effects [13].

We have recently pointed out that one important consequence of late curve divergence in cancer immunotherapy is that extended follow-up times may be required to correctly assess therapeutic efficacy [14]. Furthermore, Hoos et al. [12] have proposed a methodological framework to tackle the unique characteristics of immunotherapies. These authors consider that adherence to study designs developed for chemotherapy may have been inappropriate and contributed to failures of several immunotherapy efficacy trials. The methodological considerations put forward by Hoos et al. are highly relevant, but we would like to draw attention to two additional issues: the reason for delayed divergence of survival curves, and the importance of using adequate statistical methods in immunotherapeutic trials.

Possible reasons for delayed divergence of survival curves in immunotherapy

Delayed separation of survival curves means that the event rate is similar in treated patients and controls during the first phase after start of treatment, but that there is a relative lack of late events in treated patients compared to controls during later phases. In principle, there are two ways of explaining this phenomenon:

Delayed action: late benefit model

Earlier discussions of the reason for delayed separation of survival curves in immunotherapy trials have focused on the fact that the development of a clinically effective immune response may take time and that, therefore, the benefit of the treatment will not be evident until late, sometimes only after several months of treatment [9]. This model gains some support from patients with metastatic disease undergoing immunotherapy, who initially display progression but then later respond to therapy [15–17]. In addition, typical examples of late curve divergence that agree with the delayed action-late benefit model can be derived from studies of melanoma and prostatic cancer [2–4]. However, although a delayed clinical efficacy was observed in these studies, cellular responses have commonly been detected within weeks after vaccination, suggesting that the time to mount an immune response does not fully explain the delayed efficacy [15, 18]. Furthermore, the finding, in one study [8], that a marked delayed separation of survival curves may occur more than 3 years after end of therapy makes it unlikely that late development of immune response can be the only factor responsible for delayed separation of survival curves in immunotherapy trials.

Early action: late benefit model

Although the model of delayed action—late benefit may be operative in some cases of delayed curve divergence, an alternative, early action—late benefit scenario may be considered to explain the phenomenon. An early action of immunotherapy can be expected in any situation where there is a preexisting immune response that is boosted by the therapeutic measures utilized. Such boosting may occur in, e.g., treatment with the drug ipilimumab where the trial results have displayed a clear late and persistent curve divergence and thus can be used as prototypes of early action-late benefit model. Given the delayed clinical efficacy of ipilimumab demonstrated in Kaplan–Meier plots at the population level [1], it is tempting to draw the conclusion that the effect on the tumor is also delayed. This conclusion, however, is not warranted. The therapeutic actions leading to the lack of late events may be exerted at any time prior to the curve divergence and are thus not necessarily related in time to the graphically demonstrable effect.

During the last decade, studies have unraveled the existence of a complex and dynamic relationship between the immune system and the tumor: i.e.. the balance between tumor elimination on the one hand and tumor escape on the other hand [19, 20]. Immunotherapies aim at enhancing immune reactivity against cancer and push the escape-elimination balance toward tumor elimination. In patients with rapidly progressing disease, which may be related to high tumor load, low tumor cell immunogenicity, high proliferative or invasive capacity, etc., the task may be overwhelming and the immune system will fail to eliminate the tumor. In these cases, immunotherapy is futile, and these patients relapse and die regardless of therapeutic intervention.

In a fraction of patients with more slowly progressing disease, immunotherapy may make the difference that enables the immune system to eradicate the tumor. This immune-mediated clearance may occur early, or even very early, in treated patients, but the survival curves will still not separate until the time when corresponding control patients relapse and die. Thus, what graphically appears as a delayed effect may in fact be the result of early immunological actions leading to eradication of the tumor cells. Late divergence of survival curves is therefore an expected consequence in any case where the therapy preferentially affects tumors associated with slow disease progression. In cases of slow progression that is causally related to low tumor load, it is logical to assume that efficient elimination of tumor cells is most likely to occur at the start of treatment, when the tumor load is particularly low. Consequently, late separation of survival curves is probably often causally related to early actions of the treatment. Paradoxically, late divergence of survival curves may then, provided that the divergence persists and not replaced by curve convergence, be indicative of tumor cell eradication in certain patients, thus causing a cure of the disease.

It is important to realize that the statement that immunotherapy is preferentially active in patients with slowly progressing disease is relative and only refers to comparisons performed within the particular group of patients being studied. Thus, as amply demonstrated in many studies, also in groups of patients with rapidly progressing disease, such as metastatic melanoma or prostatic cancer [2], there are some patients who benefit from treatment and therefore do not experience the late events occurring in control patients. Interestingly, in a recently published long-term follow-up study of ipilimumab in patients with metastatic melanoma, 11 out of 15 complete responders displayed an objective clinical response within 2 months of

treatment, suggesting that early onset of action predominates in successfully treated patients even in advanced disease [21]. Conversely, as demonstrated in the study by Stadler et al. [8], patients with slowly progressing disease do not automatically benefit from immunotherapy. The reason for this is likely multifaceted and may, for instance, be explained by a relative lack of immunogenicity of the tumor cells in certain categories of patients.

Implications of late survival curve divergence

The implications of late survival curve divergence are, as we have reported previously, obvious with respect to the requirement of extended follow-up times [14]. Another important consequence of delayed separation is the implications for the statistical analysis of immunotherapy trials [22, 23]. Late divergence means that the hazard ratio is not constant over time. Standard statistical methods, such as the logrank test and the Cox proportional hazards model, are optimal to detect differences between survival curves with constant hazard ratio, but they are not ideal in immunotherapeutic trials, in which it is known or suspected a priori that late divergence of survival curves will occur [24]. To obtain a certain acceptable power in such trials, one might use, e.g., the logrank test and compensate for the non-proportional hazard by increasing the sample size, as suggested by Hoos et al. [10]. However, alternative statistical strategies to assess the clinical effect of immunotherapeutic candidates are potentially more effective and should therefore be considered [23], especially in cases where the ultimate goal is to determine long-term survival rates. Elaborate methods to handle these issues have been suggested by, e.g., Sposto et al. and Klein et al. [25, 26]. A simple and transparent strategy is to use an approximate test to compare two survival curves at a fixed late point in time, based on the Kaplan–Meier estimates and their standard errors, and the normal distribution [27]. The method of Sposto et al. tests the hypothesis of equal survival for all times after a predetermined late time T and thus also makes use of the events after this time. Readers are further referred to Logan et al. [28] for a comparison of these and several other methods, including the weighted logrank test.

Example of late curve divergence

An example of late separation of survival curves, illustrating the early action-late benefit model proposed above and showing the importance of the choice of statistical methods for the evaluation of efficacy, is provided by a study on the adjuvant treatment of cutaneous melanoma with an immunomodulatory drug, i.e., natural, multisubtype, interferon-alpha (nIFN-α) [8, 14]. In the study, adjuvant nIFN-α was given for 6 months, preceded by two cycles of dacarbazine to patients with melanoma in different stages. No significant difference with respect to overall survival rates in treated patients and controls was demonstrated at 5 years, which was the predefined time of final follow-up. However, an extended follow-up demonstrated that the survival curves gradually diverged thereafter, finally resulting in survival rates that were clearly higher in treated patients than in controls. Still, as demonstrated in Table 1, an efficacy analysis using the logrank test did not yield a statistically significant result ($p = 0.052$). By contrast, testing the equality of 9-year survival revealed a statistically significant impact on long-term survival ($p = 0.012$) [14]. With the test by Sposto et al., testing the null hypothesis that the survival curves coincide at 3 years and thereafter, the result is also significant ($p = 0.030$).

The natural course of diseases varies between different subgroups, and the results of the study described above could potentially be influenced by the heterogeneity of the patient material. Retrospective analyses of a homogenous, separately randomized, subgroup of 106 separately randomized patients with regional lymph node metastases (stage 3b) showed that late deaths were more common in controls than in treated patients. Thus, deaths occurred at almost equal rates in treated and control patients before

Table 1 Preferential effect by multisubtype IFN-α treatment on late mortality in patients with melanoma

	Average mortality (95 % CI)		Estimated 9-year OS (% ± SE)	*p* value 9-year OS	*p* value logrank
	Follow-up 0–3 years	Follow-up >3 years			
All patients, treated ($n = 128$)	12.3 (9.0–16.7)	3.2 (1.8–5.4)	57.6 ± 4.5	0.012	0.052
All patients, controls ($n = 124$)	11.5 (8.3–15.9)	9.3 (6.7–13.0)	41.7 ± 4.5		
Stage 3b, treated ($n = 54$)	21.6 (14.7–31.7)	2.2 (0.7–7.0)	45.9 ± 6.9	0.009	0.143
Stage 3b, controls ($n = 52$)	17.9 (11.8–27.2)	16.3 (10.3–25.9)	22.1 ± 6.0		

The table shows mortality per 100 person-years with 95 % confidence intervals by follow-up time interval and treatment arm, in the trial on the effect of adjuvant therapy of melanoma with natural, multisubtype IFN-α performed by Stadler et al. [8]. Early and late mortality was arbitrarily defined as occurring before and after 3 years of follow-up, respectively

3 years of follow-up, whereas after 3 years, the average mortality was about seven times higher in controls than in treated patients. With respect to statistical tests, the logrank test showed no statistically significant difference between the two groups although the 9-year OS rate was twice as high in the treated patients as in the controls, yielding $p = 0.009$ for the direct comparison. The test by Sposto et al. also yields a significant result ($p = 0.045$).

Conclusions

The reasons why Kaplan–Meier survival curves of treated patients and controls sometimes display a delayed divergence in successful immunotherapeutic trials have not been sufficiently well clarified. The phenomenon has usually been interpreted as being the result of a delayed action of the therapy, resulting in a late benefit for the patients. Although this may be true in some cases, another, more general explanation for the phenomenon can be offered, built on the simple observation that late curve divergence is always associated with a relative lack of late events among treated patients as compared to controls. Late events are indicative of slowly progressing disease, and delayed divergence will therefore be a consequence of preferential effect of treatment on tumors causing slowly progressing disease within the patient population studied. The inhibition of late events may be caused by immunological actions occurring at any time preceding the events and sometimes be a consequence of very early actions of the therapy. Therefore, early tumor cell eradication postulated to occur according to the cancer immunoediting concept may, in a seemingly paradoxical way, be manifested as clinical efficacy at a late stage, resulting in late curve divergence.

The occurrence of late curve divergence has important implications, not only for the planning of follow-up times and for the evaluation of biological markers of therapeutic response, but also for the statistical analysis of clinical trials in immunotherapy. As exemplified in this article, commonly used statistical methods, assuming proportional hazard ratios in the evaluation of efficacy, are not optimal in cases of late curve divergence. Therefore, such methods should be replaced by tests that compare survival at or after prespecified fixed late time points.

Conflict of interest The authors declare no conflict of interest.

References

1. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, Gonzalez R, Robert C, Schadendorf D, Hassel JC, Akerley W, van den Eertwegh AJ, Lutzky J, Lorigan P, Vaubel JM, Linette GP, Hogg D, Ottensmeier CH, Lebbe C, Peschel C, Quirt I, Clark JI, Wolchok JD, Weber JS, Tian J, Yellin MJ, Nichol GM, Hoos A, Urba WJ (2010) Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med 363(8):711–723. doi:10.1056/NEJMoa1003466
2. Kantoff PW, Schuetz TJ, Blumenstein BA, Glode LM, Bilhartz DL, Wyand M, Manson K, Panicali DL, Laus R, Schlom J, Dahut WL, Arlen PM, Gulley JL, Godfrey WR (2010) Overall survival analysis of a phase II randomized controlled trial of a Poxviral-based PSA-targeted immunotherapy in metastatic castration-resistant prostate cancer. J Clin Oncol 28(7):1099–1105. doi:10.1200/jco.2009.25.0597
3. Small EJ, Schellhammer PF, Higano CS, Redfern CH, Nemunaitis JJ, Valone FH, Verjee SS, Jones LA, Hershberg RM (2006) Placebo-controlled phase III trial of immunologic therapy with sipuleucel-T (APC8015) in patients with metastatic, asymptomatic hormone refractory prostate cancer. J Clin Oncol 24(19):3089–3094. doi:10.1200/jco.2005.04.5252
4. Sosman JA, Unger JM, Liu PY, Flaherty LE, Park MS, Kempf RA, Thompson JA, Terasaki PI, Sondak VK (2002) Adjuvant immunotherapy of resected, intermediate-thickness, node-negative melanoma with an allogeneic tumor vaccine: impact of HLA class I antigen expression on outcome. J Clin Oncol 20(8):2067–2075
5. Testori A, Richards J, Whitman E, Mann GB, Lutzky J, Camacho L, Parmiani G, Tosti G, Kirkwood JM, Hoos A, Yuh L, Gupta R, Srivastava PK (2008) Phase III comparison of vitespen, an autologous tumor-derived heat shock protein gp96 peptide complex vaccine, with physician's choice of treatment for stage IV melanoma: the C-100-21 Study Group. J Clin Oncol 26(6):955–962. doi:10.1200/jco.2007.11.9941
6. Wolchok JD, Neyns B, Linette G, Negrier S, Lutzky J, Thomas L, Waterfield W, Schadendorf D, Smylie M, Guthrie T Jr, Grob JJ, Chesney J, Chin K, Chen K, Hoos A, O'Day SJ, Lebbe C (2010) Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. Lancet Oncol 11(2):155–164. doi:10.1016/s1470-2045(09)70334-1
7. Wood C, Srivastava P, Bukowski R, Lacombe L, Gorelov AI, Gorelov S, Mulders P, Zielinski H, Hoos A, Teofilovici F, Isakov L, Flanigan R, Figlin R, Gupta R, Escudier B (2008) An adjuvant autologous therapeutic vaccine (HSPPC-96; vitespen) versus observation alone for patients at high risk of recurrence after nephrectomy for renal cell carcinoma: a multicentre, open-label, randomised phase III trial. Lancet 372(9633):145–154. doi:10.1016/s0140-6736(08)60697-2
8. Stadler R, Luger T, Bieber T, Kohler U, Linse R, Technau K, Schubert R, Schroth K, Vakilzadeh F, Volkenandt M, Gollnick H, Von Eick H, Thoren F, Strannegard O (2006) Long-term survival benefit after adjuvant treatment of cutaneous melanoma with dacarbazine and low dose natural interferon alpha: A controlled, randomised multicentre trial. Acta Oncol 45(4):389–399
9. Finke LH, Wentworth K, Blumenstein B, Rudolph NS, Levitsky H, Hoos A (2007) Lessons from randomized phase III studies with active cancer immunotherapies—outcomes from the 2006 meeting of the Cancer Vaccine Consortium (CVC). Vaccine 25(Suppl 2):B97–B109. doi:10.1016/j.vaccine.2007.06.067
10. Hoos A, Eggermont AM, Janetzki S, Hodi FS, Ibrahim R, Anderson A, Humphrey R, Blumenstein B, Old L, Wolchok J (2010) Improved endpoints for cancer immunotherapy trials. J Natl Cancer Inst 102(18):1388–1397. doi:10.1093/jnci/djq310
11. Hansson J (2006) Adjuvant therapy of cutaneous melanoma—current status. Acta Oncol 45(4):369–372. doi:10.1080/02841860600768895
12. Hoos A, Britten CM, Huber C, O'Donnell-Tormey J (2011) A methodological framework to enhance the clinical success of cancer immunotherapy. Nat Biotechnol 29(10):867–870. doi:10.1038/nbt.2000

13. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, Jouary T, Schadendorf D, Ribas A, O'Day SJ, Sosman JA, Kirkwood JM, Eggermont AM, Dreno B, Nolop K, Li J, Nelson B, Hou J, Lee RJ, Flaherty KT, McArthur GA (2011) Improved survival with vemurafenib in melanoma with BRAF V600E mutation. The New England journal of medicine 364(26):2507–2516. doi:10.1056/NEJMoa1103782
14. Thoren FB, Strannegard O (2011) Adjuvant interferon: extended follow-up times needed? Lancet Oncol 12(5):419
15. Berd D, Sato T, Cohn H, Maguire HC Jr, Mastrangelo MJ (2001) Treatment of metastatic melanoma with autologous, hapten-modified melanoma vaccine: regression of pulmonary metastases. Int J Cancer [Journal international du cancer] 94(4):531–539
16. Hodi FS, Butler M, Oble DA, Seiden MV, Haluska FG, Kruse A, Macrae S, Nelson M, Canning C, Lowy I, Korman A, Lautz D, Russell S, Jaklitsch MT, Ramaiya N, Chen TC, Neuberg D, Allison JP, Mihm MC, Dranoff G (2008) Immunologic and clinical effects of antibody blockade of cytotoxic T lymphocyte-associated antigen 4 in previously vaccinated cancer patients. Proc Natl Acad Sci USA 105(8):3005–3010. doi:10.1073/pnas.0712237105
17. Wolchok JD, Weber JS, Maio M, Neyns B, Harmankaya K, Chin K, Cykowski L, de Pril V, Humphrey R, Lebbe C (2013) Four-year survival rates for patients with metastatic melanoma who received ipilimumab in phase II clinical trials. Ann Oncol. doi:10.1093/annonc/mdt161
18. Kantoff PW, Higano CS, Shore ND, Berger ER, Small EJ, Penson DF, Redfern CH, Ferrari AC, Dreicer R, Sims RB, Xu Y, Frohlich MW, Schellhammer PF, Investigators IS (2010) Sipuleucel-T immunotherapy for castration-resistant prostate cancer. N Engl J Med 363(5):411–422. doi:10.1056/NEJMoa1001294
19. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD (2002) Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol 3(11):991–998
20. Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ (2011) Natural innate and adaptive immunity to cancer. Annu Rev Immunol 29:235–271. doi:10.1146/annurev-immunol-031210-101324
21. Prieto PA, Yang JC, Sherry RM, Hughes MS, Kammula US, White DE, Levy CL, Rosenberg SA, Phan GQ (2012) CTLA-4 blockade with ipilimumab: long-term follow-up of 177 patients with metastatic melanoma. Clin Cancer Res 18(7):2039–2047. doi:10.1158/1078-0432.CCR-11-1823
22. Bilusic M, Gulley JL (2012) Endpoints, patient selection, and biomarkers in the design of clinical trials for cancer vaccines. Cancer Immunol Immunother 61(1):109–117. doi:10.1007/s00262-011-1141-0
23. (2011) Guidance for industry: clinical considerations for therapeutic cancer vaccines. Department of Health and Human Services, Food and Drug Administration 76:68768–68769. https://federalregister.gov/a/2011-28726
24. Bland JM, Altman DG (2004) The logrank test. BMJ 328(7447):1073. doi:10.1136/bmj.328.7447.1073
25. Sposto R, Stablein D, Carter-Campbell S (1997) A partially grouped logrank test. Stat Med 16(6):695–704
26. Klein JP, Logan B, Harhoff M, Andersen PK (2007) Analyzing survival curves at a fixed point in time. Stat Med 26(24):4505–4519. doi:10.1002/sim.2864
27. Kalbfleisch J, Prentice R (2002) The statistical analysis of failure time data, 2nd edn. Wiley, New York
28. Logan BR, Klein JP, Zhang MJ (2008) Comparing treatments in the presence of crossing survival curves: an application to bone marrow transplantation. Biometrics 64(3):733–740. doi:10.1111/j.1541-0420.2007.00975.x