Retroductive, Comparative, Epidemiological Cohort Study with Parallel Groups Design for Evaluation of Efficacy and Safety of Drugs with ‘Well-Established Use’

Experience with the Long-Term Treatment Using the European Mistletoe Extract (Viscum album L.) in Addition to Conventional Oncological Therapy in Primary, Non-Metastatic Breast Carcinoma

P. R. Bock, W. E. Friedel, J. Hanisch, M. Karasmann, B. Schneider

Institute for Applied Medical Research, IFAG Basel AG, Basel, Switzerland
Hospital Bad Bocklet, Dept. of Oncology, Germany
Institute of Medical Statistics, University of Hannover, Medical School, Hannover, Germany

Key Words
Breast carcinoma · Mistletoe · Efficacy · Retroductive study · Epidemiological cohort study

Abstract
The randomized controlled clinical trial (RCT) is accepted as the ‘golden standard’ for the evaluation of efficacy and safety of new drugs. In contrast, to demonstrate efficacy and safety of drugs with ‘well-established use’ that have been on the European Community market for long time, observational comparative epidemiological studies can be used according to the European drug regulation directive. However, because comparative epidemiological cohort studies can share some risk of bias with other nonrandomized observational study designs, there is a need for an approach that could effectively reduce the bias risk in this type of studies. Study Objectives: The purpose of the study was to evaluate the therapeutic efficacy and safety of a long-term complementary therapy of primary, non-metastatic breast carcinoma patients treated with standardized European mistletoe extract IscadorTM (‘mistletoe’) in addition to the conventional adjuvant oncologic therapy, and compared to the control group treated with the conventional therapy alone. Methods: The multicenter, comparative, retroductive, pharmaco-epidemiological cohort study with parallel groups design and randomly selected centers that routinely used both treatments was carried out according to Good Epidemiological Practice rules under a standard operating procedure control. The test group patients received the mistletoe extract treatment subcutaneously for at least 3 months, while the control group patients of the same cohort were exclusively treated with the conventional therapy. The patients were followed up for at least 3 years or until death. The primary endpoint of efficacy was the incidence of adverse reactions to the conventional oncologic therapy. Secondary endpoints were change from baseline of the symptoms associated with the disease and treatment as well as overall survival. All endpoints were adjusted to baseline imbalance and confounders. Safety was assessed descriptively by the number of patients with adverse drug reactions (ADRs) attributed to the test treatment. Results: 1442 patients (710 tests and 732 controls) were eligible for the ‘per protocol’ analysis of efficacy and safety. At baseline, the test group had a more advanced disease and worse prognostic factors profile. After a median follow-up of 66 vs. 60 months, and a median mistletoe therapy duration of 52 months, significantly fewer test group patients (16.2%) than control patients (54.0%) developed ADRs attributed to the conventional therapy [adjusted odds ratio, OR (95% confidence interval, CI), OR = 0.47 (0.32–0.67), p < 0.001]. In the test group, the majority of the symptoms disappeared more frequently, and overall mortality hazard was significantly lower [adjusted hazard ratio, HR (95% CI), HR = 0.46 (0.22–0.96), p = 0.038] than in the control group. Systemic ADRs attributed to the test treatment developed in 0.8%, and local ADRs in 17.3% of the patients. ADR severity was mild to intermediate. Tumor enhancement was not observed. Conclusions: Complementary therapy of patients with primary, non-metastatic breast carcinoma with the mistletoe extract Iscador was safe and in comparison to the control group within the same study cohort showed con-
siderably fewer ADRs attributed to concurrent conventional therapy, reduced disease symptoms, and suggested a significant improvement of survival. Despite some methodical limitations that require careful study planning and conduction as well as critical interpretation, the applied study design seems suitable to evaluate the efficacy and safety of drugs with ‘well-established use’, particularly in oncology.

**Introduction**

Prospective randomized clinical trials (RCTs) are accepted as indispensable for proving the clinical efficacy and safety of new drugs [1–3]. RCTs are also needed to demonstrate clinical efficacy according to the requirements of Evidence-Based Medicine (EBM) rules for evidence level I [4]. In contrast, to evaluate the clinical efficacy and safety of drugs with ‘well-established use’ that were in use in the European Community (EC) for long time, according to the EC drug regulation directive the epidemiological cohort studies, particularly controlled epidemiological studies, can be used under defined preconditions [5]. Well-designed and conducted controlled epidemiological cohort studies can also satisfy the EBM requirements for evidence level II [4, 6]. However, since comparative epidemiological cohort studies might share some risk of bias with other nonrandomized observational study designs, an approach is needed that can effectively reduce the bias risk in this type of studies. To this end, an optimized approach for pharmacoepidemiological retrolective cohort studies designed for evaluation of efficacy and safety of drugs with ‘well-established use’ was developed and applied in several comparative drug treatment studies [7–9]. This approach was designed in accordance with Good Epidemiological Practice (GEP) rules [10, 11] and the IFAG standard operating procedures (SOPs) for retrolective cohort studies [12, 13]. In earlier studies, optimized comparative observational study approaches were repeatedly validated against RCTs and successfully used for evaluation of clinical therapy [14–20].

**Basic Principles of the Retrolective, Comparative, Epidemiological Cohort Method for Evaluation of Efficacy and Safety of ‘Drugs with Well-Established Use’**

The basic principles of retrolective cohort studies are internationally established within the framework of epidemiological methods. Cohort studies involve a defined population of patients followed forward in time and they evaluate the potential predictor, demography, risk, exposure, treatment, and confounding variables at baseline, and the subsequent disease outcomes in the course of the study [21–23]. A retrolective cohort study is characterized by a sampling of anonymous data from medical records in standardized case report forms (CRFs) and by a follow-up starting from the origin (i.e. diagnosis or primary surgery) in the past, with pre-specified endpoint(s) in the past, present, or future [21]. Because a retrolective cohort study is not a clinical trial but a non-interventional comparative observational study with all treatments finalized before the study commencement, some obligations inherent in GCP clinical trials, such as patients’ informed consent and insurance, ethical committee vote, test drug sample management, and approval by health authorities are not needed. Retrolective cohort studies usually have the advantage to be substantially larger, much less expensive and less time-consuming than prospective studies. However, in a retrolective comparative epidemiological cohort study, like in other nonrandomized observation studies, a special emphasis should be given to the possible risk of bias and the approaches for its sufficient reduction. The most frequent sources of bias that could occur with this type of studies are selection bias by pre-selection of patients for inclusion, exclusion of patients by selection after inclusion, patients lost for follow-up, structural homogeneity of demography and risk factor profile at baseline, nonrandom therapy differences between the groups, quality and plausibility of retroactively collected data, confounding effects by other factors (covariates), and missing values in covariates and endpoint criteria.

In the presented optimized study design we addressed the risk of bias by applying following design measures in a large representative cohort of patients:

1. Design, realization, monitoring, evaluation, interpretation, and reporting followed the GEP rules, adapted to the GCP (ICH) guidelines, and performed according to the IFAG-SOPs for retrolective cohort studies documented in a study-specific ‘operational manual’.

2. A GEP- and GCP-conforming study protocol with all study details was written before initiation of the cohort study.

3. Study centers were selected as a defined population or a representative random sample from the eligible centers of interest that accepted the study conditions and agreed to participate.

4. Patients were acquired from the pre-defined cohort population without selection, except for the pre-specified eligibility criteria according to the study protocol. All eligible patients from a center were included in chronologic order up to the pre-specified maximum number, irrespective the received treatment and outcome.

5. The medical and laboratory data records of a center were utilized as the data source that has been made anonymous by the responsible physician. If appropriate, also data from other medical centers and institutions (such as cancer registries) were used. The anonymous patients’ and therapy data were transferred to standardized case report forms (CRFs).

6. In a random sample of centers, a pre-test of comprehension, plausibility, and reliability of the CRFs and the data quality from the medical records was performed prior to the study commencement.

7. The study monitors received study-specific training. Center staff was trained in handling the CRF documentation by pre-monitoring instructions. Monitoring and data quality control were performed according to the study protocol.
8. Standardized CRF coding according to international standards (such as WHO/NIH-CTC and ICD-10) was used by professional staff. An additional CRF data check for plausibility and validity was carried out at the data entry.

9. Standardized data acquisition from the CRFs was performed by computerized data entry system with internal plausibility checks and double data entry if requested. Internal quality audit was carried out by comparison of the final electronic data base with the data from the original CRFs.

10. External data quality audit (‘source data verification’) was performed by comparison of the data base with a random sample of the original medical records and the results of the face-to-face interviews with the doctors (centers) by independent auditors.

11. As reference therapy (control group), an established (standard) treatment used in the same patient cohort was chosen, for which efficacy and safety were previously proven for the same indication in prospective randomized clinical trials. For selected indications, the control group remained without any specific treatment (i.e. ‘watchful waiting’), providing this were medically appropriate.

12. As primary endpoint(s) criteria of efficacy, usually (a) typical, objective, standardized, well-established, as well as routinely documented disease-typical event(s) or measure(s) was/were selected and evaluated according to the ‘per protocol’ principle. Alternatively, disease-typical, standardized, well-documented, summarized symptoms (such as a standardized symptom score) were used as endpoint of efficacy.

13. As the primary endpoint criterion of safety, the number and severity (CTC) of ‘adverse drug reactions’ (ADRs) or ‘unwanted events’ (UE) were used. For this analysis, the original complete study data were included according to the ‘intention-to-treat’ method.

14. The type, duration, treatment measures, and outcome of the ADRs or UE were completely documented and analyzed as secondary criteria of safety.

15. In case of baseline inhomogeneity of endpoint criteria, risk- and prognostic criteria, patients’ demography or treatment regimen, appropriate statistical methods such as multivariate adjusting and stratification were applied to substantially reduce the ‘confounding bias’. The adjusting was performed according by analysis of variance/covariance, logistic regression, or the Cox proportional hazard regression method, as well as by applying the propensity score for analysis and adjusting according to Rosenbaum, Rubin et al. [24–27].

16. The test of hypotheses concerning the primary endpoints of efficacy and safety was performed according to the study protocol by appropriate statistical methods including the estimate of effect size and its 95% confidence intervals.

17. By means of sensitivity analyses, validity, plausibility, robustness, and the ability to generalize the study results were evaluated by carrying out analyses to check reproducibility of the main evaluation, to perform planned subgroup analyses, and to calculate model analyses addressing the potential impact of excluded cases, missing values, and the choice of the covariates. For instance, the possible impact of exclusion of patients due to severe protocol violation was evaluated by comparison of the ‘intention-to-treat’ versus ‘per protocol’ analyses and the impact of missing covariate values was explored by comparison of the exclusion from analysis versus multivariate imputation.

18. The structure of the final report of the retrolective cohort study complies with the ICH-conforming reporting rules, as well as the ‘CONSORT’ recommendations [28].

**Experience with the Long-Term Treatment Using the European Mistletoe Extract (Viscum album L.) in Primary, Non-Metastatic Breast Carcinoma**

(In full length accepted for publication in Drug Research, 2004.)

**Objective:** The objective of the study was to evaluate the therapeutic efficacy and safety of the long-term post-surgical complementary treatment of primary, non-metastatic breast carcinoma patients with a standardized European mistletoe extract (‘Iscador’: Iscador®, Weleda A G, A Relsheim, Switzerland) administered in addition to conventional adjuvant onco (‘Iscador’) logic therapy with chemo-, radio- or hormonal therapy (‘conventional therapy’), compared with patients treated with conventional therapy only (control group).

**Study design:** The study was designed as a multicenter, comparative, retrospective, pharmaco-epidemiological, cohort study with parallel group design without intervention, in accordance with Good Epidemiological Practice (GEP) rules and the IFAG SOPs for retrospective cohort studies that were presented in the previous section on methods.

**Selection of centers:** The study centers were randomly selected from oncologic centers in Germany and Switzerland that have been treating patients with primary, non-metastatic breast carcinoma for several years with chemo-, radio- and hormonal therapy with or without additional Iscador treatment, provided they accepted the study protocol, including monitoring and data audit, as well as agreed to participate in the study.

**Selection of patients:** In each center, all eligible patients were included in the study in chronologic order, without any further selection and irrespective of the study course or disease outcome, only limited by the pre-specified maximum sample size which was set to a total of 1600 patients.

**Eligibility and exclusion criteria:** All completely documented patients’ records were eligible for the study irrespective of the disease outcome and treatment compliance, providing they were treated following surgery between 1988 and 2000 for at least 6 months in the selected centers using the conventional therapy with or without complementary Iscador treatment and followed up for at least 3 years, or until death. Exclusion criteria were any severe protocol violations such as recurrent or metastatic breast carcinoma at primary surgery, or treatment with other mistletoe products, as well as incomplete cases with missing essential data. Patients with pre-
mature study termination for any reason were not eliminated but followed up to the last available data values.

Treatment: Treatment allocation to a particular conventional therapy regimen with or without Iscador was carried out according to the individual patients' health status and preference at the discretion of the treating physicians. It was finished before the study commencement, hence not influenced by the study. Eligible patients were either treated with 2–3 weekly subcutaneous injections of Iscador in addition to the conventional therapy (test group), or received the conventional therapy alone (control group).

Endpoint criteria: The primary endpoint criterion of efficacy was the rate of adverse drug reactions (ADRs) attributed to the conventional therapy. The secondary endpoints of efficacy were the rate of complete symptoms associated with disease and treatment ('symptoms') and the survival of the patients during the study and follow-up. Survival analysis was performed by calculating the hazard ratio for the overall survival (OS) outcomes adjusted for baseline imbalance, treatment regimen and confounder effects with the Cox proportional hazard regression method. The evaluation of efficacy was performed according to the 'per protocol' approach. The possible impact of the exclusion on the results was evaluated in a sensitivity analysis. The primary endpoint of Iscador safety was assessed by the rate of systemic and local ADRs explicitly attributed to the Iscador treatment. Type, number, severity and outcome of all Iscador-attributed ADRs were evaluated according to WHO/NIH/CTC-criteria [29, 30]. The safety results were confirmed by a sensitivity analysis of all originally acquired cohort patients, including those with severe protocol violation ('intention-to-treat' approach).

Demographic and prognostic baseline criteria: Age, general and disease-specific medical history, body weight, hormone receptor status, menopause, date, type and outcome of the surgical treatment, present symptoms, concurrent diseases and treatments, duration of the present disease as well as any previous therapy were documented as baseline criteria. Baseline assessment of disease severity and prognostic profile consisted of tumor stage (TNM, UICC), grade, localization, multiplicity at surgery, type of surgical treatment such as resection versus ablation, with or without axillary dissection, and the surgery result expressed as tumor-free status (NED) or residual tumor, as well as the initial symptoms and the Karnofsky index.

Data acquisition and quality assurance: Clinical investigators, who were instructed and supervised by professional monitors according to the GEP and GCP rules, transferred the anonymous medical records to 165 (10.3%) were excluded due to severe protocol violation, from the total of 1607 originally acquired patient data records 165 (10.3%) were excluded due to severe protocol violation, such as pre-existing recurrence or metastatic disease, or missing essential data, leaving 1442 (710 test and 732 control) eligible patients (mean age 53 and 57 years) for evaluation of efficacy and safety according to the 'per protocol' approach. A sensitivity analysis did not detect any significant bias on the adjusted outcome results due to the exclusion.

Patient baseline characteristics: The baseline values showed significantly more severe and advanced disease in the test group (Table 1). The initial symptoms, particularly gastrointestinal, CNS, skin/mucosal symptoms as well as recurrent infections and pain were significantly more frequent in the test group. Median follow-up duration (66 vs. 60 months) as well as the mean time from surgery to treatment initiation (1.4 vs. 1.2 months) did not relevantly differ between the treatments.

Treatment regimens: During the study, 43.9% (test group) vs. 75.7% (control group) of patients were treated with at least one course of radiotherapy, while 32.8% vs. 23.2% received at least one course of chemotherapy. Hormonal therapy was similar for both groups (Table 2). The test group patients received Iscador therapy for a median of 52 months. On average, Iscador injections were administered subcutaneously 2–3 times per week.

Reduction of ADRs associated with the conventional therapy: Significantly less test group patients (16.2%) than control patients (54.0%, p < 0.001) suffered from ADRs attributed to conventional therapy such as skin and mucosal reactions, including mucositis, gastrointestinal- and CNS symptoms. The adjusted risk estimate for experiencing an ADR attributed to the conventional therapy was expressed as adjusted OR = 0.47 (0.32–0.67), p < 0.001. This result is indicating a remarkable independent protective effect of the complementary mistletoe treatment with Iscador against the risk of ADRs attributed to the radio-, chemo-, or hormonal therapy in early-stage breast carcinoma patients. The significant bene-
Table 1. Baseline characteristics

<table>
<thead>
<tr>
<th>Baseline demographic and prognostic criteria (initial sample size 710 vs. 732)</th>
<th>Test group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years (mean)</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>Body weight, kg (mean)</td>
<td>67</td>
<td>72</td>
</tr>
<tr>
<td>Diagnosis-to-surgery time, months (mean)</td>
<td>0.36</td>
<td>0.14</td>
</tr>
<tr>
<td>Estrogen receptor positive, %</td>
<td>72.5</td>
<td>65.1</td>
</tr>
<tr>
<td>Post-menopause, %</td>
<td>61.6</td>
<td>82.4</td>
</tr>
<tr>
<td>Tumor stage ‘high-risk’ (pT2–4), %</td>
<td>61.4</td>
<td>50.3</td>
</tr>
<tr>
<td>Tumor stage node positive (N &gt; 0), %</td>
<td>46.3</td>
<td>40.7</td>
</tr>
<tr>
<td>Tumor grade ‘high-risk’ (G3–4), %</td>
<td>44.9</td>
<td>21.2</td>
</tr>
<tr>
<td>Tumor multicellular, %</td>
<td>25.0</td>
<td>8.9</td>
</tr>
<tr>
<td>Any (&gt;1) surgical interventions, %</td>
<td>7.7</td>
<td>1.6</td>
</tr>
<tr>
<td>Residual tumor after surgery, %</td>
<td>2.7</td>
<td>0.5</td>
</tr>
<tr>
<td>A llergy in the medical history, %</td>
<td>12.8</td>
<td>7.6</td>
</tr>
<tr>
<td>Other concurrent (non-oncologic) diseases, %</td>
<td>36.0</td>
<td>53.9</td>
</tr>
</tbody>
</table>

‘test’ = Test group received Iscador treatment in addition to the conventional therapy; ‘control’ = control group received conventional therapy only; ‘conventional therapy’ = radio- and/or chemo- and/or hormonal therapy.

Table 2. Overview of treatment regimens and supportive therapy

<table>
<thead>
<tr>
<th>Treatment regimen (initial sample size = 710 vs. 732)</th>
<th>Test group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiotherapy received, % (including any combination)</td>
<td>43.9</td>
<td>75.7</td>
</tr>
<tr>
<td>Chemotherapy received, % (including any combination)</td>
<td>32.8</td>
<td>23.2</td>
</tr>
<tr>
<td>Hormonal therapy received, % (including any combination)</td>
<td>50.1</td>
<td>50.3</td>
</tr>
<tr>
<td>Median Iscador: therapy duration, months</td>
<td>52</td>
<td>not applicable</td>
</tr>
<tr>
<td>Median cumulative Iscador dose, mg</td>
<td>4354</td>
<td>not applicable</td>
</tr>
<tr>
<td>Median study and follow-up duration, months</td>
<td>66</td>
<td>60</td>
</tr>
<tr>
<td>Mean time from surgery to begin of therapy, months</td>
<td>1.36</td>
<td>1.21</td>
</tr>
<tr>
<td>Concurrent antiemetic drug treatment, %</td>
<td>8.6</td>
<td>4.9</td>
</tr>
<tr>
<td>Concurrent analgesic drug treatment, %</td>
<td>4.5</td>
<td>6.0</td>
</tr>
<tr>
<td>Concurrent physical, or balneotherapy, %</td>
<td>18.9</td>
<td>35.1</td>
</tr>
<tr>
<td>Concurrent RDA vitamin treatment, %</td>
<td>24.8</td>
<td>7.9</td>
</tr>
</tbody>
</table>

‘test’ = Test group received Iscador treatment in addition to the conventional therapy; ‘control’ = control group received conventional therapy only; ‘conventional therapy’ = radio- and/or chemo- and/or hormonal therapy; ‘mono’ = monotherapy applying the particular treatment; RDA = recommended daily allowance for vitamins, minerals and trace elements.

The presented comparative retrolective cohort study with 1442 evaluable primary, non-metastatic breast carcinoma patients shows a substantial statistically significant reduction of ADRs attributed to adjuvant conventional therapy, a significantly higher rate of symptom relief, as well as a significant improvement of the overall survival in the Iscador group when compared with the control group without Iscador. These results, particularly the effect on survival, however, need to be interpreted with some caution because the applied study design shares some potential risk for bias with other nonrandomized observational studies. We attempted to effectively minimize these potential biases by utilizing a large study cohort from randomized centers that used treatment regimens either with or without Iscador, by applying a standardized parallel...
group study design with strict adherence to the study protocol, by application of GEP rules and SOPs with standardized and blinded data documentation in CRFs, by integrated control, monitoring, independent auditing of data quality, and by multivariable adjusting of endpoint criteria for baseline imbalance, treatment regimen and other potential confounders. Sufficient quality of the standardized data acquisition from the medical records was reported earlier [31]. This optimized comparative retrospective cohort study method was already successfully used in the clinical evaluation of other therapeutic regimens [7–9].

It seems very unlikely that the treatment effect on the endpoint criteria could be biased in favor of Iscador by the baseline imbalance between the treatment groups, not only due to the effective adjusting for confounders, but also because initially more patients of the test group showed severe symptomatic disease and a worse prognostic factor profile. Further, the results from numerous sensitivity analyses with the present study under various model conditions consistently and sufficiently reproduced the results of the main analysis and hence did not indicate the presence of any effective hidden confounder.

With regards to the quality-of-life criteria as well as the symptoms associated with the disease and therapy, the published evidence of a substantial clinical benefit from a complementary mistletoe treatment in cancer patients appears convincing and conclusive. The results of the present study suggest a relevant and significant benefit from Iscador treatment concerning symptoms and quality of life in early-stage breast cancer patients. This is supported by earlier controlled studies, reporting a significant benefit on quality of life as well as on disease- and treatment-associated symptoms from mistletoe treatment in patients with various solid cancers (e.g. [32, 33]). In a recent multicenter, comparative, retrospective cohort study a standardized mistletoe therapy in primary, non-metastatic breast carcinoma was associated with significantly fewer ADRs from conventional therapy and showed a significant improvement in symptoms associated with disease and treatment when compared with controls without mistletoe therapy [9]. Regarding the toxicity, the Iscador treatment was well tolerated without any life-threatening ADRs, particularly without severe allergic reactions. The rate of systemic ADRs and local reactions at the injection site had about the same magnitude and quality as previously published from other clinical trials on mistletoe therapy [34–36]. Tumor enhancement was not observed. Consequently, complementary Iscador treatment in breast carcinoma patients can be regarded as safe.

In conclusion, despite some possible methodological limitations inherent to any non-randomized design, the results of the present study suggest convincing evidence for a significant and clinically relevant benefit from the complementary Iscador therapy regarding the incidence of the ADRs attributed to the conventional oncologic therapy, the improvement of the disease- and treatment-associated symptoms and quality of life, and also suggest a relevant and significant improvement of the overall survival, particularly in patients with properly conducted post-surgical conventional oncologic therapy and a long-term Iscador treatment. The Iscador-treatment was well tolerated and can be regarded as safe.

References