Bayesian statistics and clinical trial designs for human cells and tissue products for regulatory approval

Yoji Jokura, Kazuo Yano, Natsumi Watanabe, Masayuki Yamato

Abstract

Introduction: In order to obtain premarket approval for medical products derived from human cells or tissues in the United States (US), the European Union (EU), and Japan, data from clinical trials are typically required to evaluate product efficacy and safety. Clinical investigators or study sponsors often face challenges when designing clinical trials on human cells and tissue products with the goal of obtaining premarket approval owing to the unique characteristics of products in this category. The methods used to administer, infuse and transplant these products vary more widely than the methods used for pharmaceuticals. In addition, final product quality may vary depending on the product source, i.e., patients or donors. These products are generally intended to treat intractable and rare diseases or injuries; therefore, it may not be possible to collect a sufficient number of cases and enrollment may be a long process. Moreover, since the technology for product development in this category is relatively new, knowledge and experience from previous studies are limited.

Methods: The key elements for the design of clinical trials to determine product efficacy were identified by examining clinical trial designs for approving products. Review reports for approved products from regulatory authorities in the US and Japan as well as the European public assessment reports in the EU were analyzed.

Results: For one product approved in the US, Dermagraft®, Bayesian statistics were used to evaluate product efficacy, instead of traditional (frequentist) statistics. Based on the statistical guidance for clinical trials recently issued by the US Food and Drug Administration, statistical analyses including Bayesian statistics are key elements in the design of clinical trials for products based on human cells and tissues. New regulations regarding human cells and tissue products have recently been implemented in Japan, including conditional and time-limited approval for regenerative medicine products. In these cases, Bayesian statistics are a promising alternative approach to support product development.

Conclusions: Our results emphasize the benefit of considering cogitating statistical methods, such as Bayesian statistics, when designing clinical trials for regulatory purposes.

1. Introduction

Medical products derived from human cells and tissues are unambiguously distinct from chemically synthesized drugs and medical devices, and are regulated separately by the authorities; they are categorized as “human cells, tissues, and cellular and tissue-based products (HCT/Ps)” in the United States (US) [1], “advanced therapy medicinal products (ATMPs)” in the European Union (EU) [2], and “regenerative medicine products” in Japan [3,4].
When clinical investigators or study sponsors conduct clinical trials on human cells and tissue products with the goal of obtaining premarket approval, they often face challenges, particularly when assessing product efficacy, owing to unique characteristics of products in the category. There is wider variation in the methods used to administer, infuse and transplant these products than in those for pharmaceuticals. In addition, the quality of the final product may vary depending on the source, i.e., patients or donors. These products are generally intended to treat intractable and rare diseases or injuries; therefore, it may not be possible to collect a sufficient number of cases and study enrollment may be very slow. These features of cells and tissue products that may impact clinical study design were summarized in a guidance of the US Food and Drug Administration (FDA) [5]. Moreover, since the technology used for products in the category is relatively new, knowledge and experience from previous studies are limited.

In an analysis of the study designs of clinical trials on human cells and tissue products approved in the US, the EU, and Japan, we found that Dermagraft®, which utilizes allogeneic cells, was approved in the US based on pivotal study data that was analyzed using Bayesian statistics. In this pivotal study, an interim analysis, which is considered an adaptive design, was utilized, and these statistical approaches were important for the acceptance of the clinical efficacy of the product.

Bayesian statistics are an alternative to traditional statistics, i.e., frequentist statistics, and have recently been employed in clinical trials to evaluate pharmaceuticals, not only in Phase III studies, but also in Phase I and II studies [6]. In some cases, Bayesian statistics can be used to reduce the sample size and to apply mid-course adjustments to a trial design, or to stop a trial, shortening the study duration [7]. Moreover, Bayesian statistics have been regarded as a useful statistical method for clinical trials since the middle of the last decade because the approach is ideally suited to adapting to information accrued during a trial, potentially allowing for smaller and more informative trials [8]. In the pivotal Dermagraft® study, an interim analysis was utilized in the decision to stop the study when a targeted number of cases was reached, and to determine the necessity for additional enrollment. The purpose of an interim analysis is to stop a trial early if a sufficient difference between groups is obtained to conclude that an intervention is effective or harmful [9]. Early stopping may allow subjects in the placebo arm as well as those not in the trial to receive a beneficial treatment sooner. In contrast, when severe side effects are encountered, early stopping of the trial may prevent unnecessary harm. Early stopping may also save money and facilitate the rapid reporting and translation of results to clinical practice [9].

The FDA recently issued a guidance to the study design of clinical trials for medical devices, named “Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials” [7]. Although some additional conditions are required to utilize Bayesian statistics: 1) prior information should be discussed with the FDA prior to the initiation of a study, and 2) indications of the device may be impacted by modifications at the interim analysis, the Bayesian framework has several unique advantages over its frequentist counterpart. Traditional statistical methods only use information from previous studies at the design stage. In contrast, Bayesian statistics formally incorporate prior information gathered before, during, and outside of the trial [10]. Furthermore, many clinical trials are conducted over an extended period of time, and it is desirable to frequently monitor the interim results of such trials in order to promote more rapid decisions when sufficient evidence is obtained. Bayesian methods allow for more frequent monitoring and interim decision making during trials [10]. Based on the concept outlined in the FDA guidance as well as the successful example, Bayesian statistics should be considered in the design of studies to evaluate human cells and tissue products.

No such guidance documents for statistical methods used in clinical trials have been introduced by regulatory authorities in the EU and Japan; this initiative is unique to the US. Statistical inferences are based on mathematical models of experiments, including clinical trials. Moreover, in Japan, new regulations for human cells and tissue products were introduced in 2014, in which conditional and time-limited approval pathways specific to the product category are included [4].

In the current study, we performed a comparative investigation of the guidance documents for statistical methods for clinical trials in the US, the EU, and Japan, and summarized of a unique case in which Bayesian statistics and an interim analysis were successfully applied during a trial design.

2. Methods

Guidance documents describing statistical methods for clinical trials were obtained from appropriate regulatory websites in the US [11], the EU [12], and Japan [13]. Approval information for human cells and tissue products was obtained from the websites of the relevant regulatory authorities in the US (Biologics [14], Premarket approval (PMA) [15], Humanitarian Device Exemption (HDE) [16]), the EU [17], and Japan [18] at the end of June, 2016. According to the definitions and research methods used in previous studies [19,20], products utilizing either autologous or allogeneic human cells or tissues were selected from review reports in the US and Japan and from European public assessment reports in the EU. The study design for each product was identified from the clinical data section in each report. Individual review reports or European public assessment reports of the products approved in the US, the EU, and Japan were obtained from the following sources: the FDA websites for Carticel™ [21], Epicel® [22], Provenge® [23], Laviv® [24], Dermagraft-TCTM® [25], Apri-grafTM/Garftskin [26], Composite Cultured Skin [27], OrcelTM [28], Dermagraft® [29], Gintuit® [30], Hemacord [31], HPC/Cord blood [32], Ducord [33], Allocord [34], HPC/Cord blood [35], and HPC cord blood [36]; European Medical Agency (EMA) websites for ChondroCelect® [37], MACI® [38], Provenge [39], and Holoclar [40]; and Pharmaceuticals and Medical Devices Agency (PMDA) websites for JACC [41], JACE [42], Temcell™ HS Inj. [43] and Heart-Sheet® [44]. Information related to clinical trials for the approved products, such as that described at ClinicalTrials.gov [45], was also analyzed.

3. Results

3.1. Guidance on clinical trial design

In the US, several guidance documents related to clinical trial design for medical devices and human cells and tissue products have recently been issued by the FDA, “Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials” was issued by the Center for Devices and Radiological Health (CDRH) and the Center for Biologics Evaluation and Research (CBER) of the FDA on February 5, 2010 [7]. This guidance document clarifies the utilization of Bayesian statistics, and includes a strong recommendation for consultation with the FDA when planning a study protocol. It is specifically applicable to medical device clinical trials, including products derived from human cells and tissues. CDRH and CBER also issued the draft guidance document “Adaptive Designs for Medical Device Clinical Studies” on May 18, 2015 [46]. This document addresses adaptive designs for medical device clinical trials and is applicable to pre-market medical device submission,
including PMA, premarket notification (510(k)) submissions, de novo submissions, Humanitarian Device Exemption (HDE) applications, and Investigational Device Exemption (IDE) submissions [46]. More recently, CBER of the FDA issued a guidance document for industry titled “Considerations for the Design of Early-Phase Clinical Trials of Cellular and Gene Therapy Products” in June 2015 [5]. The guidance document was issued to assist sponsors and investigators with the design of early-phase clinical trials for cellular therapy and gene therapy products, and to provide current recommendations regarding clinical trials in which the primary objectives are the initial assessment (most Phase I and some Phase II) of safety, tolerability, or feasibility of the administration of investigational products [5].

In the EU, “Statistical Principles for Clinical Trials” was published by the EMA [47]. These guidelines were developed at the International Conference on the Harmonization of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), and apply to clinical trials on “medicinal products” or drugs. In Japan, the Ministry of Health and Welfare issued guidelines as a notification titled “Statistical Principles for Clinical Trials” [48], translated from the ICH guidelines introduced in the EU described above. The following statement describes the application of Bayesian statistics in the ICH guidelines: “…because the predominant approaches to the design and analysis of clinical trials have been based on frequentist statistical methods, the guidance largely refers to the use of frequentist methods when discussing hypothesis testing and/or confidence intervals. This should not be taken to imply that other approaches are not appropriate: the use of Bayesian and other approaches may be considered when the reasons for their use are clear and when the resulting conclusions are sufficiently robust.”

3.2. Approval and clinical study design for human cells and tissue products and associated clinical designs

According to the definitions presented in previous studies [19,20], 24 products derived from human cells or tissues for clinical treatments were identified as of the end of June 2016. Clinical trial information (e.g., study design, sample size, primary efficacy endpoints, and results) obtained for products is summarized in Table 1. Six products, Hemacord® [31], HPC/Cord blood [32], Ducord® [33], Allocord® [34], HPC/Cord blood [35], and HPC cord blood [36] were excluded owing to a lack of clinical studies or trials that solely involved the target product; hence, information for these products is not available in the review reports. When we analyzed the review reports for products in the US and Japan as well as the European public assessment reports for products in the EU, we found that a reasonably large number of cases were typically enrolled and primary endpoints were evaluated statistically, except for the humanitarian device exemption (HDE) products in the US and most approved products in Japan. As described in Section 3.3, product information for Dermagraft® was investigated in great detail because the clinical study design was unique from a regulatory standpoint owing to its use of an interim analysis and Bayesian statistics. The study was initiated in December 1998 and completed in March 2000, and the study results, which were evaluated in the review by the FDA, were published in 2003 [49].

3.3. Dermagraft®

Dermagraft® is a cryopreserved human fibroblast-derived dermal substitute; it is composed of fibroblasts, an extracellular matrix, and a bioabsorbable scaffold. The product is manufactured from human fibroblast cells derived from newborn foreskin tissue. The application of the product was reviewed by CDRH and approved in September 28, 2001 [29].

Dermagraft® was intended for use in the treatment of full-thickness diabetic foot ulcers with a duration of greater than six weeks and those that extend through the dermis, but without tendon, muscle, joint capsule, or bone exposure. Dermagraft® is recommended for use in conjunction with standard wound care regimens and in patients that have adequate blood supply to the involved foot.

The sponsor conducted a total of five trials to evaluate the product, and the data obtained in all trials were used for assurance of the safety and efficacy of the product in the treatment of full-thickness diabetic foot ulcers with a duration of greater than six weeks by product review. Pivotal studies were conducted twice to evaluate product efficacy, Pivotal I and II (Table 1), and data from the second pivotal study, Pivotal II, were used to confirm the product efficacy in the authority review for product approval. In both pivotal studies, the superiority of Dermagraft® with the conventional treatment over the conventional treatment alone as the control was evaluated. The primary efficacy endpoint for both pivotal studies was the same, i.e., complete wound closure by 12 weeks. In the first pivotal study, or Pivotal I, the primary endpoint was challenging for comparisons with the control treatment, as defined in the protocol; however, no conclusion was reached regarding the superiority of the product with respect to the control treatment in the review report, even though “42 out of the 109 evaluable patients from the Dermagraft® group (39%) and 40 out of the 126 evaluable patients from the control group (32%) reached complete wound closure by 12 weeks.”

The second pivotal study, or Pivotal II, had a similar study design to Pivotal I (i.e., a randomized control study in which patients were treated with Dermagraft® plus conventional therapy, or conventional therapy alone). The completion of wound closure by 12 weeks was evaluated as the primary efficacy endpoint, as in Pivotal I. In this trial, the original total sample size was calculated based on a two-group Fisher’s exact test with a 0.0294 one-sided significance level and at least 80% power, requiring an enrollment of up to 330 patients for the Dermagraft® and control groups (1:1) in the initial study design (Fig. 1) [49]. The original statistical plan called for an interim analysis to be performed after 180 patients completed the study. The 0.0294 level of significance (using the Pocock method [50]), which guided the Data Monitoring Committee with respect to study continuation, was not achieved in the interim analysis [49]. In the interim analysis, the relationship between ulcer duration at the time of screening and the incidence of ulcer healing with Dermagraft® was analyzed based on the results of a modified statistical plan, which indicated that (1) the efficacy analysis should be based only on patients with ulcers with a duration of greater than six weeks at the time of the screening visit, and (2) the primary endpoint should be analyzed using Bayesian statistics. Furthermore, information obtained during the initial part of the trial (the interim analysis) was utilized prospectively in the latter part in order to estimate overall efficacy (in the final analysis). Additional patients were enrolled until the required endpoint of the Bayesian sequential procedure was achieved (98.4% probability of benefit) [49]. A diagram of the required patient numbers and actual numbers enrolled in the trial is shown in Fig. 2. Although 314 patients were enrolled in the study, only 245 patients exhibited ulcers with a duration of greater than six weeks were included in the final analysis. Based on the Bayesian analysis, the probability that Dermagraft® plus conventional therapy increased the chance of achieving wound closure in patients with ulcers with a duration of greater than six weeks over that of conventional therapy alone was 98.4%. Although the original sample size calculated using Fisher’s exact
Table 1
Summary of pivotal study designs for evaluating the efficacy of human cells and tissue products.

| Product name (Category, approval date, authorities) | Indication | Study/Design | Sample size | Primary endpoint | Results of the primary endpoint |
|-----------------------------------------------------|------------|--------------|-------------|------------------|---------------------------------|
| **US**                                              |            |              |             |                  |                                 |
| Dermagraft-Tc™ (PMA, March 18, 1997, FDA/CDRH)      | For use as a temporary wound covering for surgically excised full-thickness and deep partial-thickness thermal burn wounds in patients prior to autograft placement | Randomized, controlled, within-patient, unmasked | 66 (within-patient control) | % Autograft “take” at 14 days. (Comparison with standard care: cryopreserved cadaver allograft) | Significantly equivalent to that of wounds treated with allografts (94.7% for Dermagraft-Tc™ vs. 93.1% for frozen cadaver allografts (control), p = 0.0001) |
| Carticel™ (BLA, August 22, 1997, FDA/CBER)           | For the repair of clinically significant, symptomatic, and cartilaginous defects of the femoral condyle (medical, lateral, or troclear) caused by acute or repetitive trauma | 1. Swedish retrospective clinical study: | 153 (consecutive patients) | Clinical outcome by the question, “how does your knee feel now compared to before surgery?” | Patient questionnaire: 70% reported an improved status |
|                                                      |            | 2) Retrospective case report forms | 2) Questionnaire | 191 (completion of 12-month follow-up: 38) |                                 |
|                                                      |            | 3) Biopsy date (n = 25) | 3) Patient questionnaire |                                 |                                 |
|                                                      |            | 2. U.S. registry data base | 4. U.S. registry data base |                  |                                 |
|                                                      |            |              |             |                  |                                 |
| Apligraf™/Graftskin (PMA, May 28, 1998, FDA/CDRH)    | For use with standard therapeutic compression for the treatment of non-infected partial and full-thickness skin ulcers due to venous insufficiency of greater than 1 month duration and which have not adequately responded to conventional ulcer therapy | Prospective, randomized, controlled, multi-specialty, unmasked | 161 (Apligraf™), 136 (Control) | 1) The incidence of 100% wound closure per unit time, and 2) The overall incidence of 100% wound closure by 6 months | 1) 50% patients achieved wound closure: 140 days (Apligraf™) and 181 days (Control), (p = 0.3916) 2) 55.4% (72/130, Apligraf™) and 49.1% (54/110, Control). (p = 0.365) |
| Composite Cultured Skin (HDE, February 21, 2001, FDA/CDRH) | For use in patients with mitten hand deformity due to Recessive Dystrophic Epidermolysis Bullosa (RDEB) as an adjunct to standard autograft procedures for covering wounds and donor sites created after the surgical release of hand contractions | 1. Australian clinical study: with-in patient historical control | 7 (historical control: 5) | Duration of digital functionality | The use of CCS and autografts did not decrease the time to re-surgery |
|                                                      |            | 2. United States clinical study: | 12 (within-patient control) | Healing in Composite Cultured Skin (CCS) treatment of donor sites | When CCS was used, the need for donor sites was reduced. In addition, CCS-treated donor sites healed without complications |
|                                                      |            | within-patient controlled, randomized | 2) Collagen sponge | The incidence or time to wound healing in comparison to CCS, the acellular sponge and standard non-adherent dressing at any time point | No significant differences |
|                                                      |            | 3) Standard care | 3) Patient questionnaire | The time (days) to wound closure (100% re-epithelialization) |                                 |
|                                                      |            |              |             | 82 |                                 |
| **Europe**                                          |            |              |             |                  |                                 |
| Orcel™ (PMA, August 31, 2001, FDA/CDRH)              | For the treatment of fresh, clean split thickness donor site wounds in burnt patients | Multicenter, randomized, within-patient controlled | 139 (Dermagraft®) | Number of patients reached complete wound closure by 12 weeks Complete wound closure by 12 weeks | 15 days for Orcel™ vs. 22 days for the Control, for the ITT population, median days to 100% wound closure (p = 0.0006, Log-Rank test) 30% (42/139, Dermagraft®) and 28% (40/142, control) in ITT analysis |
|                                                      |            |              |             | 142 (Control) |                                 |
|                                                      |            |              |             | 163 (Dermagraft®) |                                 |
|                                                      |            |              |             | 151 (Control) |                                 |
|                                                      |            |              |             | 130 (Dermagraft®) |                                 |
|                                                      |            |              |             | 115 (Control) |                                 |
| Dermagraft® (PMA, September 28, 2001, FDA/CDRH)      | For use for the treatment of full-thickness diabetic foot ulcers of greater than six weeks duration which extend though the dermis, but without tendon, muscle, joint capsule, or bone exposure | 1. Randomized, controlled, masked (Pivotal I) | 20 (Standard Care plus Epicel™), 24 (Standard care only) | Mortality | Mortality*: 10.0% for Epicel™ vs. 62.0% for Standard Care only *Significant in the original article |
|                                                      |            |              |             | |                                 |
|                                                      |            | 2. Randomized, controlled, masked (Pivotal II) | | | |
| Epicel™ (HDE, October 25, 2007, FDA/CDRH)           | For use in patients who have deep dermal or full-thickness burns comprising a total body surface area of greater than or equal to 30% | Physician-sponsored study: prospective, single-centered, controlled, randomized | | | |
|                                                      |            |              |             | | |

(continued on next page)
| Product name (Category, approval date, authorities) | Indication | Study/Design | Sample size | Primary endpoint | Results of the primary endpoint |
|---------------------------------------------------|------------|--------------|-------------|-----------------|----------------------------------|
| **Provenge**® (BLA, April 29, 2010, FDA/CBER)     | For the treatment of asymptomatic or minimally symptomatic metastatic castrate-resistant (hormone refractory) prostate cancer | Phase 3: randomized, double-blind, placebo-controlled | 341 (Provenge®) 171 (Placebo) | Overall survival | Median survival (months) Sipuleucel-T (Provenge®): 25.8 Placebo: 21.7 (p = 0.032) |
| **Laviv**® (BLA, June 21, 2011, FDA/CBER)          | For improvements of the appearance of moderate to severe nasolabial fold wrinkles in adults | Two Phase 3, multicenter, double blind, controlled | 1. IT-R-005 100 (Laviv®) 103 (Vehicle) 2. IT-R-006 110 (Laviv®) 108 (Placebo) | 1) Two-point improvement in the Evaluator Wrinkle Severity Assessment (6-point scale) 2) Two-point improvement in the Subject Wrinkle Assessment (5-point scale) | - Subject assessment: 57% for Laviv® vs. 30% for Control, (p = 0.0001)  - Physician Assessment: 33% for Laviv® vs. 7% for Control, (p = 0.0001)  - Subject assessment: 45% for Laviv® vs. 18% for Control, (p < 0.0001)  - Physician Assessment: 19% for Laviv® vs. 7% for Control, (p = 0.0075) |
| **Gintuit** (BLA, March 9, 2012, FDA/CBER)         | For topical (non-submerged) application to a surgically created vascular wound bed in the treatment of mucogingival conditions in adults. Gintuit is not intended to provide root coverage | Prospective, randomized, within-subject controlled | 96 | Efficacy: 85 | Percentage of Gintuit sites with KT* ≥2 mm at six months, compared to a 50% success rate, in a single-arm comparison * keratinized tissue 95.3% met success criteria at Gintuit, (p < 0.001)**: Comparison to a pre-defined standard of 50% of subjects with KT width ≥2 mm |
| **EU** ChondroCelect® (ATMP, October 5, 2009, EMA/CHMP) | For use in the repair of single symptomatic cartilage defects of the femoral condyle of the knee (International Cartilage Repair Society [ICRS] grade III or IV) in adults | Phase 3, multicenter, randomized, controlled | 57 (ChondroCelect®) 61 (Microfracture (Active comparator)) | Superiority on the structural repair (histology) endpoint at 12 months and non-inferiority on the clinical endpoint (change from baseline in KOOS) for the average of the 12- to 18-month follow-up data | Differences in the endpoint:  - Histomorphometric 0.26 (p = 0.003)  - ICRSII at 12 months 10.92 (p = 0.0103) Change in KOOS at 12 and 18 months 1.81 (no significant differences) Difference LS (least squares) means:  - Pain 11.76 (P < 0.001)  - Function 11.41 (P < 0.001) |
| **MACI** (ATMP, June 27, 2013, EMA/CHMP)           | For use in the repair of symptomatic, full-thickness cartilage defects of the knee (grade III and IV of the Modified Outerbridge Scale) of 3–20 cm² in skeletally mature adult patients | Prospective, randomized, open-label, parallel-group, multicenter | 72 (MACI) 72 (Microfracture (Active comparator)) | Change from baseline to Week 104 for the patient’s Knee injury and KOOS pain and function (Sports and Recreational Activities [SRA]) scores | Difference LS (least squares) means:  - Pain 11.76 (P < 0.001)  - Function 11.41 (P < 0.001) |
| **Provenge** (ATMP, September 6, 2013, EMA/CHMP)   | For the treatment of asymptomatic or minimally symptomatic metastatic (non-visceral) castrate-resistant prostate cancer in male adults in whom chemotherapy is not yet clinically indicated | Randomized, double-blind, controlled, multicenter | 341 (Provenge) 171 (Placebo) | Overall survival | Median survival (months) Provenge: 25.8 Placebo: 21.7 (p = 0.032) |
| **Holoclar** (ATMP, February 17, 2015, EMA/CHMP)   | For the treatment of adult patients with a moderate to severe limbal stem cell deficiency (defined by the presence of superficial corneal neovascularization in at least two corneal quadrants, with central corneal involvement, and severely impaired visual acuity), unilateral or bilateral, due to physical or chemical ocular burns | Retrospective, non-randomized, uncontrolled, case series-based observational | 104 | ACLSCT success (success of transplantation) | ACLSCT success (rate %): 75 (72.1%) (p < 0.001), 95% CI (62.5, 86.5) |
| Japan | JACE (Medical device, October 29, 2007, MHLW/PMDA) | For use in patients with serious, extensive burns when sufficient donor sites for autologous skin graft are not available and the total area of deep dermal and full-thickness burns is 30% or the total surface area | Multi-center, open-label, uncontrolled | 2 | Formation of epidermis at 4 weeks, scoring from 1 to 4 | Very effective: 1 (50%) Effective: 1 (50%) |
|------|-------------------------------------------------|----------------------------------------------------------------------------------|-----------------------------------|---|-----------------------------------------------|----------------------------------------|
| JACC (Medical device, July 27, 2012, MHLW/PMDA) | To alleviate clinical symptoms of the traumatic cartilage deficiency and osteochondritis dissecans (excluding knee osteoarthritis) in the knee joints with a cartilage-defective area of 4 cm² or more for which there are no other options | Prospective, multi-center, non-randomized, uncontrolled | 30 | Composite endpoint with functional improvement in the knee and arthroscopic evaluation (ICRS score) | Very effective: 25 (83.3%) Effective: 3 (10%) Neither: 2 (6%) Not effective: 0 (0.0%) |
| Temcell® HS Inj. (Regenerative medicine product, September 18, 2015, MHLW/PMDA) | For the treatment of acute graft-versus-host disease (GVHD) following hematopoietic stem cell transplant | Single-arm, open-label, phase 2-3 | 25 | Complete response continued for 28 days post treatment | 48% (12/25 cases, 95% CI: 27.8, 68.7) |
| HeartSheet® (Regenerative medicine product, September 18, 2015, MHLW/PMDA) | For use in treatment of severe HF (heart failure) caused by ischemic heart disease, despite maximal standard-of-care drug and interventional therapies meeting satisfying all conditions: - New York Heart Association (NYHA): class III or IV - A left ventricular ejection fraction (LVEF) ≤ 35% on resting | Single-arm, open-label, phase 2 | 7 | The change in LVEF on gated equilibrium blood-pool scintigraphy from pre-transplantation to 26 weeks post-transplantation | Improved: 0 Unchanged: 5 Worsened: 2 |

ACLSCT, Autologous Cultured Limbal Stem Cell Transplantation; ATMP, Advanced Therapy Medical Products; BLA, Biologics License Application; CBER, Center of Biologics Evaluation and Research; CDRH, Center of Device and Radiological Health; CHMP, Committee for Human Medicinal Products; EMA, European Medicines Agency; EU, European Union; FDA, Food and Drug Administration; HDE, the Humanitarian Device Exemption; ICRS, International Cartilage Repair Society; IDE, Investigational Device Exemption; ITT, Intention-To-Treat; KOOS, Knee injury and Osteoarthritis Outcome Score; LVEF, left ventricular ejection fraction; MHLW, Ministry of Health, Labor and Welfare; PMDA, Pharmaceuticals and Medical Devices Agency; US, the United States.

* The following six cord blood products are also approved in the US, but were excluded from the study owing to a lack of clinical studies or trials that solely examined the product: Hemacord (BLA, Nov. 10, 2011, CBER), HPC Cord blood (BLA, May 24, 2012, CBER), Ducord (BLA, Oct. 3, 2012, CBER), Allocord (BLA, May 30, 2013, CBER), HPC Cord blood (BLA, June 13, 2013, CBER), and HPC cord blood (BLA, January 28, 2016, CBER).
test (i.e., frequentist statistics) was 330, the final number of patients enrolled in the study was 314.

The authority concluded that the study provided reasonable evidence for the efficacy of Dermagraft® in the treatment of full-thickness diabetic foot ulcers with a duration of greater than six weeks, in addition to the assurance of safety from all studies as follows: the probability that Dermagraft® plus conventional therapy provided a treatment benefit to patients with ulcers with a duration of greater than six weeks that provided by conventional therapy alone was 98.4%. Furthermore, there was a 95% probability that the percentage of patients achieving wound closure with Dermagraft® ranged between 22% and 38% and that the percentage...
of patients achieving wound closure with the control treatment ranged between 12% and 26%. The inclusion criteria for efficacy endpoints were patients with diabetic ulcers for 6 weeks or longer at the time of screening. Thus, the approved indication based on the results of the clinical trial was changed to patients with diabetic ulcers for 6 weeks or longer.

4. Discussion

As of the end of June 2016, 24 products were approved as human cells and tissue products in the US, the EU, and Japan. The study designs and product characteristics, such as the size of the target population, indications, approval category (e.g., HDE), source of the materials, and clinical expectations (e.g., performance), varied markedly, and these properties may affect study designs. Among human cells and tissue products approved under current regulations in the US, some early products were approved based on only clinical registry data in the US and with only a small number of study cases. In Japan, three of four products were approved based on a pilot or phase I/II studies lacking a hypothesis, and with a very small number of cases from clinical trials, as shown in Table 1. In order to support product development in the industry and to provide clinicians access to new human cells and tissue products, the regulatory authorities in the US, the EU, and Japan have taken various measures, such as establishing regulations and providing guidance. In the US, several guidance documents have been published with recommendations regarding statistical designs for clinical trials of human cells and tissue products. In Japan, new initiatives, including new regulations for human cells and tissue products, were implemented in 2014; two products, i.e., Temcell® HS Inj. and Heart-Sheet®, were approved in September 2015 as regenerative medicine products under these new regulations [51]. Furthermore, HeartSheet® was approved under a conditional and time-limited approval pathway, which is another new initiative dedicated to the regenerative medicine product category under the new regulations [52].

In our analysis of efficacy evaluations, we found that the pivotal study for Dermagraft® used a Bayesian statistical analysis to evaluate the product by incorporating the results of an interim analysis, and the authority judged the product efficacy based on these results. In the review of Dermagraft®, “prior information” from the interim analysis was used for the final analysis, which was described as follows: Bayesian statistics allow for information obtained during the initial part of a trial to be utilized prospectively in the latter part in order to enable the overall estimation of measures of efficacy [29]. In addition, considering that frequentist statistics were used for the first pivotal study and Bayesian statistics were used in the second pivotal study for product approval, it is beneficial to consider a flexible trial design.

Bayesian statistics are suitable for clinical trials to evaluate products, such as medical devices, because they are continuously improved; “prior information” is necessary for the analysis and is generally obtained during product development. Pennello et al. [53] of the Division of Biostatistics of the FDA explained why device trials are particularly well-suited to Bayesian analyses. “For example, if a therapeutic device has evolved in relatively small increments from previous generation of the same increments from previous generations of the same type of device, then prior information from the trials for the previous device can be predicted of the safety and effectiveness profile of the new derive. The reason the previous trials can be predictive is that the mechanism of action of a therapeutic device is often physical, implying a local effect that is often predictable. In contrast, the mechanism of action of pharmaceuticals is pharmacokinetic/pharmacodynamics, implying systemic effect that are often unpredictable form similar but not identical formulations.” Cells and tissue products that only affect or impact local tissues and do not exert systemic effects, e.g., skin substitutes, muscle substitutes, or cartilage substitutes, share similar properties to medical devices; accordingly, we concluded that prior information obtained from previous studies can be applied to new studies in a Bayesian framework. Dermagraft® is human fibroblast-derived dermal substitute for the treatment of diabetic foot ulcers, and it has primarily local effects (at the wound site) [49]. This case study supports the feasibility of a Bayesian approach for human cells and tissue products with similar properties to those of medical devices.

Methods that use information accrued during trials are adaptable, but they have various additional benefits: accumulating results may provide a basis for modifying the design of the trial, e.g., by slowing (or stopping) or expanding accrual, altering randomization strategies to favor better-performing therapies, dropping or adding treatment arms, and changing the trial population to focus on patient subsets that are responding better to the experimental therapies [8]. In addition to prior information obtained from the interim analysis of the trial, “historical” information from the source, such as clinical trials conducted overseas, patient registries, clinical data obtained for very similar products, and pilot studies, may be used according to the “Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials” [7].

The FDA issued the guidance document on February 5, 2010 that apply not only to medical device applications, but also to human cells and tissue products approved as medical devices, such as Dermagraft®. The objective of this guidance is to provide the least burdensome way to evaluate submissions by the authority. The guidance indicates that “the Bayesian approach, when correctly employed, may be less burdensome than a frequentist approach. Section 513(a) (3) of the Federal Food, Drug, and Cosmetic Act (FDCA) mandates that the FDA shall consider the least burdensome appropriate means of evaluating efficacy of a device that would have a reasonable likelihood of resulting in approval” [7]. Between issuances of the draft and the final guidance, the FDA held a workshop in 2005 [54] and public meetings between 2006 and 2009 [55] to ensure that public input is reflected in the guidance. In the period between 1999 and 2003, prior to the issuance of the draft guidance in 2006, at least 14 original PMA and PMA Supplements were approved by the FDA using a Bayesian primary analysis [56]. Moreover, in the workshop in 2005, the FDA report titled “Innovation or Stagnation: Challenges and Opportunities on the Critical Path to Medical Product Development” was published, which emphasized the need for intensified scientific efforts to improve the development and evaluation of medical products [54], resulting in the issuance of the guidance.

Two main points should be considered when applying Bayesian statistics: 1) according to the guidance, prior information should be discussed with the FDA prior to the initiation of a study, and 2) indications of the device may be impacted by modifications at the interim analysis. The former point describes a regulatory pathway in the US for consultation with the FDA when designing clinical studies, similar to the procedure for an IDE. In Japan, there is also a consultation system with regulatory authority, PMDA, which is used when designing clinical trials for regulatory approval purposes [4]. Therefore, prior to finalizing a study design, it is necessary to obtain alignments from regulatory authorities in the US and Japan. The second consideration can be broadly applied, particularly to studies using Bayesian statistics, if the results from an interim analysis are used in the final analysis. For example, if an interim analysis limits the original indications, the final indication may also be limited, as demonstrated in the case of Dermagraft®.

Bayesian statistics can be also useful for post-marketing studies, since pre-marketing study results (information) could be used as
prior information for post-marketing studies. FDA Bayesian guidance also states the following [7]: “FDA believes the Bayesian approach is well suited for surveillance purposes. The key concept: “Today’s posterior is tomorrow’s prior” allows you to use the posterior distribution from a pre-market study as a prior distribution for surveillance purpose, to the extent that data from the clinical study reflect how the device is used after approval. You may continue to update post-market information via Bayes’ theorem as more data a gathered.”

Overall, the FDA guidance indicates that a sound Bayesian approach could be less burdensome than a traditional statistical approach [7]. However, some important lessons can be gained from previous experiences. For example, Bayesian trials need to be prospectively designed. It is never a good idea to switch from a frequentist to Bayesian approach, or vice versa. Applicants need to meet early and often with regulators, as noted above. Prior information needs to be identified in advance. The control group cannot be used as a source of prior information for a new product, especially if the objective is to show that the new product is not inferior. The applicant needs to work with an expert Bayesian statistician. Computing power is necessary for large-scale simulations. Validation and Quality Assurance are key components of the study operation.

In the EU and Japan, although there were no specific guidance documents for statistical methods, there was broader guidance for clinical trials on medicinal products in which Bayesian statistics were introduced. Since the final pivotal study (Pivotal II) of Dermagraft® was conducted between 1998 and 2000 [57], the study was designed before the issuance of the guidance, and even before the draft guidance was issued in 2006 [58]. Moreover, because the guidance was only issued in 2010, it will be necessary to conduct further analyses to determine whether this guidance is useful for clinical investigations on human cells and tissue products and to obtain product approval. However, based on a literature review of clinical trials using Bayesian statistics published through September 2011, most (67.2%) of the 122 studies utilized Bayesian statistics for efficacy evaluations [10], suggesting that Bayesian statistics are appropriate for confirming product efficacy. Therefore, it is reasonable to consider Bayesian statistics when designing clinical investigations, depending on the characteristics and indication of the product, owing to its benefits, such as a potentially shorter study duration and smaller enrollment.

In addition to the statistical guidance, the FDA recently issued a draft guidance regarding adaptive designs for clinical studies, “Adaptive Designs of Medical Devices Clinical Studies,” which supports the development of human cells and tissue products based on the design of clinical investigations by “reducing resource requirements and/or increasing the chance of study success” [46]. Considering the recent issuance of this guidance and draft guidance for new statistical approaches in clinical studies, it may become a powerful tool for designing clinical trials, and it has the potential to accelerate the development of human cells and tissue products. Considering guidance related to clinical trial statistics, such as “Considerations for the Design of Early-Phase Clinical Trials of Cellular and Gene Therapy Products” published in June 2015 [5], the FDA aggressively supports sponsors or manufacturers for the development of medical products, rather than other authorities.

In Japan, a new Regulatory Act named the “Pharmaceuticals and Medical Devices Act” (PMD Act) was implemented in 2014, and a new category was established, i.e., “regenerative medicine products.” In addition to drugs, medical devices, quasi-drugs, and cosmetics [44]. Regenerative medicine products are defined as products based on processed human or animal cells intended for either a) the reconstruction, repair, or formation of a structure or its function in the human (or animal) body (i.e., tissue-engineered products) or b) the treatment or prevention of human (or animal) diseases (i.e., cell therapy products), or articles intended for the treatment of diseases in humans (or animals) that are transgenic for expression in human (or animal) cells (i.e., gene therapy products). Specifically for the new category, the PMD Act introduced conditional and time-limited approvals. If safety is confirmed and a probable benefit of the product is demonstrated by clinical trial(s), the product receives conditional and time-limited approval. After obtaining this approval, it is possible to market products; however, patient follow-ups must be conducted in order to further confirm product safety and efficacy, and data must be submitted to the authority for final approval [4]. In view of this new regulation, Bayesian statistics may be useful. Specifically, premarket trial data can be incorporated as prior information in a postmarket follow-up study to examine effectiveness for the final application for production and distribution after obtaining conditional and time-limited approval.

5. Study limitations

The survey focused on human cells and tissue products currently approved in the US, EU, and Japan. The results of this analysis are therefore limited by the small sample size of approved products compared to the numbers of drugs and medical devices.

6. Conclusion

The use of a modern statistical approach, such as Bayesian statistics, and interim analyses (i.e., an adaptive design) may be considered when evaluating the efficacy of human cells and tissue products in clinical trials, despite the difficulties associated with these approaches. Interim analyses are beneficial for identifying a target patient population for a product. Similarly, new regulations enabling conditional and time-limited approval in Japan promote the use of Bayesian statistics in follow-up studies.

Conflict of interest

Mr. Yoji Jokura is an employee of Cook Japan Inc. Dr. Kazuo Yano is an employee of Medtronic Japan Co., Ltd. Dr. Masayuki Yamato is a shareholder of CellSeed Inc.

Acknowledgments

The authors would like to thank Dr. Scott Snyder (Cook Research Inc.) for his advice on statistical descriptions.

References

[1] US Food and Drug Administration. Human cells, tissues, and cellular and tissue based products, 21CFR. 2012. Available from: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcr/CFSResearch.cfm?CFRPart=1271.
[2] European Parliament and the Council of the European Union. REGULATION (EC) No 1394/2007 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 November 2007 on advanced therapy medicinal products and amending Directive 2001/83/EC and Regulation (EC) No 726/2004. 2007. Available from: http://ec.europa.eu/health/files/eudralex/vol-1/reg_2004_726/reg_2004_726_ en.pdf.
[3] Government of Japan. The Pharmaceutical and Medical Devices Act (the act on securing quality, efficacy and safety of pharmaceuticals, medical devices, etc., Act No. 145. [In Japanese]. 1960. Available from: http://www.hourei.mhlw.go.jp/menu-ji-binsi-do/c3frame/cp/1MODE=hourei&DMODE=CONTENTS&SMODE=NORMAL&KEYWORD=AndEFSNo=635.
[4] Azuma K. Regulatory landscape of regenerative medicine in Japan. Curr Stem Cell Rep 2015;1:118–28.
[5] US Food and Drug Administration. Guidance for industry, considerations for the design of early-phase clinical trials of cellular and gene therapy products. 2015. Available from: http://www.fda.gov/downloads/Biologics.../UCM359073.pdf#research=Guidance...for-Industry&ZC%3AConsiderations...for-the-Design...of-EarlyPhase...Clinical...Trials...of-Cellular...and-%3AGene...Therapy...Products...
