A method for establishing class III medical device equivalence: sodium hyaluronate (GenVisc 850) for the treatment of knee osteoarthritis

Gheorghe Doros¹
Philip T Lavin²
Michael Daley³
Larry E Miller⁴

¹School of Public Health, Boston University, Boston, ²OrthogenRx Inc., Doylestown, PA, ³Miller Scientific Consulting, Inc., Asheville, NC, USA

Abstract: Although the concept of equivalence for drugs (generics) and biologics (biosimilars) has been readily adopted, the concept of equivalence or indistinguishable characteristics for class III medical devices has yet to be specifically addressed regarding a defined regulatory approval process in the US. In September 2015, GenVisc 850® (sodium hyaluronate), a hyaluronic acid approved for the treatment of knee osteoarthritis, was approved by the US Food and Drug Administration (FDA) based upon indistinguishable characteristics in comparison to an approved branded hyaluronic acid (Supartz®/Supartz FX™). The purpose of this paper is to review the methodology and report the main outcomes used to demonstrate clinical comparability of GenVisc 850 with Supartz/Supartz FX. The FDA approval was collectively attained using prospectively defined methods for preclinical, physical, and chemical testing, as well as noninferiority in clinical performance comparisons. Evidence from five randomized controlled studies of Supartz/Supartz FX vs saline control injections (used for Supartz approval), two randomized controlled trials of GenVisc 850 vs saline control injections, and one randomized controlled study of GenVisc 850 vs Supartz/Supartz FX provided evidence of safety for GenVisc 850. Efficacy was further assessed based on assessment of the same Supartz studies and three prospectively identified GenVisc 850 studies. A Bayesian network meta-analysis was used to demonstrate that the clinical efficacy of GenVisc 850 was noninferior to Supartz/Supartz FX and superior to saline control. Overall, safety of GenVisc 850 was similar to that of Supartz/Supartz FX and saline control injections, while efficacy of GenVisc 850 was noninferior to that of Supartz/Supartz FX and superior to saline control injections.

Keywords: Food and Drug Administration, generic, substantially equivalent, biosimilar, indistinguishable, hyaluronic acid, knee, class III, medical device, osteoarthritis

Introduction

Since the enactment of the Drug Price Competition and Patent Term Restoration Act of 1984,¹ >10,000 generic drugs have entered US market. Generic drugs now account for 88% of prescriptions dispensed in the US² with generic substitution responsible for >$239 billion in annual health-care savings.³ More recently, the Biologics Price Competition and Innovation Act of 2009,⁴ enacted as part of health-care reform (Affordable Care Act), defined an abbreviated licensure pathway for biological products shown to be biosimilar to a Food and Drug Administration (FDA)-licensed reference product. A biological product is a virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, or analogous product intended to prevent, treat, or cure a medical condition. The biosimilar pathway allows a manufacturer to demonstrate equivalence to a
branded or reference product based on establishing an identical mechanism of action, route of administration, composition, dosage form and strength, and indication for use. In March 2015 utilizing this guidance, the FDA approved filgrastim-sndz (Sandoz International GmbH, Holzkirchen, Germany), an injectable granulocyte colony-stimulating factor, as the first generic biosimilar in US pharmaceutical history. According to an analysis by the Congressional Budget Office, the continued availability of new biosimilars is projected to save $25 billion in health-care costs over the next 10 years.6

For medical devices, the 510(k) route – reserved for class II devices (moderate-to-high risk) – was introduced in 1976 to establish a pathway for determining substantial equivalence of a device to a predicate device already on the market. Since inception, the FDA has cleared >140,000 devices via the 510(k) program, allowing a pathway to commercialization for substantially equivalent class II medical devices and allowing for significant competition in the market. To date, a similar pathway for class III (high risk) medical device approval has yet to be adopted by the FDA, although a viable regulatory pathway was established ~20 years ago through the FDA Modernization Act.7

Intra-articular injections of hyaluronic acid (HA), commonly referred to as viscosupplementation, have been classified in the US by the FDA as class III medical devices for >20 years based on the presumption that the primary mechanism of action in alleviating pain is due to viscoelastic or mechanical properties of HA in solution rather than biological activity, with eleven products receiving FDA approval for the treatment of knee osteoarthritis (OA) since 1997, including Supartz®/Supartz FX™ (sodium hyaluronate; Bioventus LLC, Durham, NC, USA). In 1995, GenVisc 850® (sodium hyaluronate; OrthogenRx, Inc., Doylestown, PA, USA) was approved as a generic drug to Supartz/Artz/Artzal in Japan and is currently approved in >60 countries with extensive postmarketing safety data in >35 million doses administered to patients. The active components of GenVisc 850 (known under the branded name of Adant® outside the US) were specifically formulated and demonstrated to be equivalent in chemical composition, molecular characteristics, and dosage form to Supartz/Supartz FX. In September 2015, GenVisc 850 was approved by US FDA as a class III medical device based upon indistinguishable physical, chemical, and clinical (noninferiority) performance compared with the approved product, Supartz/Supartz FX, as well as documented superior performance compared with saline control injections. GenVisc 850 also demonstrated no unexpected safety risks, including demonstrated safety in the largest and longest duration saline-controlled clinical study of HA for knee OA.8 The purpose of this paper is to review the prospectively defined methodology, analyses, and outcomes that were utilized to demonstrate similar clinical performance and attain US marketing approval for GenVisc 850.

**Characteristics of GenVisc 850**

GenVisc 850 is a viscoelastic nonpyrogenic solution of purified, high-molecular-weight sodium hyaluronate with a pH of 6.8–7.8. The sodium hyaluronate is a polysaccharide containing repeating disaccharide units of glucuronic acid and N-acetylglucosamine derived from a bacterial fermentation process. The average molecular weight is 850 kDa with a range of 620–1,170 kDa (hence the designation GenVisc 850). Each 2.5 mL solution contains 10 mg/mL sodium hyaluronate dissolved in 1% physiological saline. GenVisc 850 is indicated for the treatment of knee OA pain in patients who have failed to respond adequately to conservative nonpharmacological therapy and simple analgesics. The characteristics and indications for use of GenVisc 850 are indistinguishable from those of Supartz/Supartz FX.9

**Preclinical and analytical evaluation**

Preclinical studies were conducted with GenVisc 850 to evaluate analytical chemistry; rheological and molecular weight dispersion; bioequivalence; acute, subacute, and chronic toxicity; mutagenicity/genotoxicity; immunogenicity/sensitization; irritation; hemolysis; release specifications; and sterilization validation. Overall, these tests confirmed that the chemical attributes, analytical characteristics, and rheological (molecular chain interactions) and molecular weight dispersion characteristics of multiple lots of GenVisc 850 were indistinguishable from various lots of the approved medical device, Supartz/Supartz FX. As part of this testing, 15 blinded samples of batches of GenVisc 850 or Supartz/Supartz FX were submitted to a testing laboratory, analyses were conducted under Good Laboratory Practices, and the samples were determined to be indistinguishable.10 Furthermore, it was observed that the iniproduct variation was not different than the interproduct variation, using these same parameters.

**Clinical standards for comparability evaluation**

**GenVisc 850 safety evaluation**

**Safety methods**

Given that GenVisc 850 and Supartz/Supartz FX were formulated to have comparable attributes that were confirmed by
analytical testing, the primary question arises as to whether such testing can translate to comparable clinical performance. The comprehensive evidence supporting safety of GenVisc 850 is summarized in the clinical results published in both the GenVisc 850 and the Supartz/Supartz FX Summary of Safety and Effectiveness Data,9,11 where clinical safety was evaluated in five prospective, randomized, double-blind studies of Supartz/Supartz FX (n=619) vs saline control (n=537) injections.11 Additional evidence of a comparable safety profile for GenVisc 850 was provided by a multicenter, double-blind, randomized trial that compared GenVisc 850 (n=116) with Supartz/Supartz FX (n=113) in patients with knee OA.12

Safety results
In the five studies of Supartz/Supartz FX vs saline control, the most common adverse events (AEs) were arthralgia, arthropathy, back pain, nonspecific pain, injection site reaction, headache, and injection site pain (Table 1). There were no statistically significant differences in the incidence of AEs between Supartz/Supartz FX and control groups. This integrated analysis demonstrated that, in relation to saline control, Supartz/Supartz FX had a comparable safety profile.

In the multicenter, double-blind, randomized trial comparing GenVisc 850 with Supartz/Supartz FX, the incidence of AEs was 18.1% vs 23.0% (P=0.36) for any AE, 1.7% vs 5.3% (P=0.17) for device-related AEs, and 0% vs 0.9% (P=1.0) for serious AEs with GenVisc 850 and Supartz/Supartz FX, respectively, after 6 weeks (Figure 1). The serious AE reported in the Supartz/Supartz FX group was prostatic hyperplasia treated with surgical excision, which was deemed unrelated to the device.

Table 1 Adverse events reported in five randomized controlled trials with Supartz/Supartz FX vs saline control injections

<table>
<thead>
<tr>
<th>Adverse event</th>
<th>Supartz/Supartz FX (n=619)</th>
<th>Saline control (n=537)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthralgia</td>
<td>110 17.8</td>
<td>95 17.7</td>
</tr>
<tr>
<td>Arthroplasty/arthrosis/arthritis</td>
<td>68 11.0</td>
<td>57 10.6</td>
</tr>
<tr>
<td>Back pain</td>
<td>40 6.5</td>
<td>26 4.8</td>
</tr>
<tr>
<td>Pain (nonspecific)</td>
<td>37 6.0</td>
<td>26 4.8</td>
</tr>
<tr>
<td>Injection site reaction</td>
<td>35 5.7</td>
<td>18 3.4</td>
</tr>
<tr>
<td>Headache</td>
<td>27 4.4</td>
<td>23 4.3</td>
</tr>
<tr>
<td>Injection site pain</td>
<td>26 4.2</td>
<td>22 4.1</td>
</tr>
</tbody>
</table>

Notes: Data extracted from Food and Drug Administration.11 Adverse events with incidence >4% are reported. Includes application/injection site reaction, injection site inflammation, and purpura at injection site.

Safety conclusion
The totality of evidence from five randomized controlled trials of Supartz/Supartz FX vs saline control injections and a randomized controlled trial of GenVisc 850 vs Supartz/Supartz FX provided adequate evidence of safety for FDA approval of GenVisc 850. Safety was further supported by the postmarketing safety reports of Supartz/Supartz FX over the past 20 years.

GenVisc 850 clinical efficacy evaluation

Efficacy methods
Clinical efficacy of the GenVisc 850 was evaluated using several separate comparative methodologies and approaches. The results of a randomized controlled trial and Bayesian meta-analysis of all available relevant studies were used to show that the clinical performance of GenVisc 850 is similar to Supartz and superior to PBS. A randomized controlled, multicenter clinical trial compared GenVisc 850 with Supartz, while Bayesian analyses evaluated the noninferiority of GenVisc 850 using data from randomized controlled trials that studied GenVisc 850 and Supartz. A first Bayesian analysis used only data from four studies that collected Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) A pain score.16 A second analysis added Karlsson study18, which used WOMAC A as outcome, however, included patients with a Kellgren–Lawrence score of 4. Additional analyses were carried out on data from nine studies (five that collected WOMAC A pain score and four that did not) and the second analysis. In these analyses, if studies had WOMAC A it was used, if not then the Lequesne score (if available) was converted into WOMAC A by multiplying by a factor of 7 to the mean change from baseline and the standard deviation (SD).

A prospectively defined Bayesian network meta-analysis using random effects was utilized for combining direct
and indirect evidence using noninformative prior distributions.\textsuperscript{13,14} Posterior inferences were carried out by sampling from the posterior distribution of the parameters using Gibbs sampling in OpenBUGS,\textsuperscript{15} where 50,000 samples are obtained from the posterior distribution after the initial 5,000 are discarded. To decrease autocorrelation, only every tenth sample was used in calculations. Mean, SD, and 95% credible interval were obtained for each marginal posterior distribution. Convergence of the Markov chain Monte Carlo sampler was checked using the Brooks–Gelman–Rubin method. Convergence of the sampler was established if the ratio of within-chain and between-chain variability for multiple chains starting at different initial values was close to 1. With our specifications, this was established after the 5,000 burn-in samples.

We hypothesized that the efficacy of GenVisc 850 and Supartz/Supartz FX was each superior to saline control and that GenVisc 850 would be noninferior to Supartz/Supartz FX. Since Supartz/Supartz FX was already approved by the FDA, documenting the superiority over saline control using this methodology would serve to validate the approach. A priori, a 4 mm difference was established as the noninferiority margin comparing GenVisc 850 with Supartz/Supartz FX.

**Efficacy results**

In the five prospective, double-blind, randomized trials of Supartz/Supartz FX vs saline control, there was a statistically greater reduction in total Lequesne Index\textsuperscript{16} score of 2.7 for Supartz/Supartz FX vs 2.0 for saline control ($P=0.003$), thereby confirming the approval conclusions of the FDA in considering the Premarket Approval (PMA) application for Supartz/Supartz FX and supporting the statistical approach in this Bayesian analysis. Furthermore, in an analysis of a randomized controlled trial of GenVisc 850 vs Supartz/Supartz FX, mean reduction in pain on movement >6 weeks assessed with a 100 mm visual analog scale was 50±21 in each group, while mean improvement in total Lequesne Index score was 5.7±3.2 for GenVisc 850 and 6.1±3.2 for Supartz/Supartz FX. While both HAs resulted in statistically and clinically meaningful improvements in knee OA symptom severity, there were no observed between-group differences.

Baseline patient characteristics were comparable among all studies (Table 2).

A first analysis included four randomized controlled studies that used WOMAC A as a primary end point, evaluating the efficacy of Supartz/Supartz FX or GenVisc 850 in patients with knee OA. This analysis excluded Karlsson study\textsuperscript{18}, which was the only one to enroll patients with a Kellgren–Lawrence score of 4. Of the four studies, two were Supartz/Supartz FX vs saline control studies, selected from the five pivotal studies used for regulatory approval, and were included based upon prospective selection criteria that are based upon the availability of WOMAC pain scores\textsuperscript{17–19} and two were GenVisc 850 vs saline control\textsuperscript{8–10} studies.

Using a Bayesian longitudinal network meta-analysis with a linear trend of studies using WOMAC A as a primary end point and with longer follow-up duration for comparing longitudinally GenVisc 850, Supartz/Supartz FX, and saline control yielded comparable reductions in WOMAC pain scores at 6 weeks, continued reductions in knee pain for >26 weeks in both HA groups, while improvements in knee pain gradually diminished in the saline control group over the same period (Figure 2). The posterior probability of superiority of GenVisc 850 vs saline control was 79% at

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatments</th>
<th>n</th>
<th>Male (%)</th>
<th>Age (years)</th>
<th>BMI (kg/m(^2))</th>
<th>K-L grade</th>
<th>OA duration (years)</th>
<th>WOMAC pain score</th>
<th>Follow-up (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blanco et al\textsuperscript{13}</td>
<td>GenVisc 850:saline</td>
<td>22:20</td>
<td>19:19</td>
<td>68:68</td>
<td>33:34</td>
<td>4.0:4.0</td>
<td>10.0:11.0</td>
<td>63:68</td>
<td>26</td>
</tr>
<tr>
<td>Navarro-Sarabía et al\textsuperscript{14}</td>
<td>GenVisc 850:saline</td>
<td>153:153</td>
<td>16:16</td>
<td>63:64</td>
<td>29:29</td>
<td>2.3:2.3</td>
<td>3.5:3.5</td>
<td>56:57</td>
<td>30</td>
</tr>
<tr>
<td>Day et al\textsuperscript{17}</td>
<td>Supartz/Supartz FX:saline</td>
<td>108:115</td>
<td>44:38</td>
<td>62:63</td>
<td>30:29</td>
<td>2.5:2.5</td>
<td>4.7:4.7</td>
<td>66:68</td>
<td>18</td>
</tr>
</tbody>
</table>

**Abbreviations:** BMI, body mass index; K-L, Kellgren–Lawrence; OA, osteoarthritis; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

![Figure 2](image)
Table 3 WOMAC pain score posterior mean differences with GenVisc 850, Supartz/Supartz FX, and saline control from randomized controlled trials included in Bayesian longitudinal meta-analysis with a linear trend

<table>
<thead>
<tr>
<th>Week</th>
<th>GenVisc 850 vs saline control</th>
<th></th>
<th>Supartz/Supartz FX vs saline control</th>
<th></th>
<th>GenVisc 850 vs Supartz/Supartz FX</th>
<th></th>
<th>Mean difference</th>
<th>Posterior probability (%)</th>
<th>Mean difference</th>
<th>Posterior probability (%)</th>
<th>Mean difference</th>
<th>Posterior probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.0</td>
<td>42</td>
<td>−1.3</td>
<td>65</td>
<td>2.3</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.7</td>
<td>44</td>
<td>−1.7</td>
<td>70</td>
<td>2.4</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>−0.6</td>
<td>60</td>
<td>−3.2</td>
<td>83</td>
<td>2.6</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>−1.5</td>
<td>69</td>
<td>−4.3</td>
<td>84</td>
<td>2.8</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>−1.8</td>
<td>71</td>
<td>−4.6</td>
<td>85</td>
<td>2.8</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>−2.8</td>
<td>75</td>
<td>−5.8</td>
<td>85</td>
<td>3.0</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>−3.1</td>
<td>75</td>
<td>−6.1</td>
<td>85</td>
<td>3.0</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>−3.7</td>
<td>76</td>
<td>−6.9</td>
<td>86</td>
<td>3.1</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>−5.6</td>
<td>78</td>
<td>−9.1</td>
<td>86</td>
<td>3.5</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>−6.9</td>
<td>79</td>
<td>−10.6</td>
<td>86</td>
<td>3.7</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

30 weeks, with a 6.9 mm mean advantage (Table 3). Additionally, the difference in WOMAC pain scores between the GenVisc 850 and Supartz/Supartz FX remained below the prespecified noninferiority margin of 4 mm throughout the 30-week follow-up period (Figure 3). Paucity of data toward the end of the time interval causes an increase in variance, and therefore, the posterior probability of noninferiority does not increase. The posterior mean difference between the effect of GenVisc 850 and the reference product was always below the noninferiority margin with a posterior probability of 50%, but the scarcity of data limits the ability to declare noninferiority of GenVisc 850 to the reference product for the interval extending to 30 weeks.

Efficacy conclusion

Results from two randomized trials comparing Supartz/Supartz FX with saline control, two randomized controlled trials comparing GenVisc 850 with saline, a randomized controlled trial comparing the GenVisc 850 with Supartz/Supartz FX, and a Bayesian meta-analysis integrating efficacy data from published randomized controlled trials all confirm that the clinical efficacy of GenVisc 850 was noninferior to Supartz/Supartz FX and both products were superior to saline control.

Discussion

This report summarizes the clinical evidence that supported a PMA (class III medical device) approval of GenVisc 850 based upon indistinguishable analytical characteristics and noninferior clinical performance to an approved branded product, Supartz/Supartz FX. GenVisc 850 was approved for the treatment of knee OA in September 2015 and is currently marketed in the US, based on indistinguishable physical and chemical characteristics and noninferior clinical performance compared with Supartz/Supartz FX, a previously FDA-approved HA.9

Although the term “generic” is reserved for drugs, and equivalent terminology has been adopted for biologics (biosimilars) and class II medical devices (substantially equivalent), a term has yet to be applied to an approval of a class III medical device. Establishing a defined regulatory pathway for approval of equivalent or indistinguishable devices, specifically for the HA class and class III medical devices, may significantly help to reduce the clinical and economic burden of knee OA. Approximately 18 million adults in the US suffer from symptomatic knee OA, with one in four reporting difficulty with ambulation, despite repeated attempts with traditional nonsurgical therapies.22,23 With the continued aging of the population, the economic burden of knee OA is projected to increase by almost 50% by 2025.23
Recently, intra-articular HA treatment was reported to be among the most effective, long-lasting, and safest treatments for OA knee pain.24–27 Recent projections suggest that the HA market may exceed $1 billion in 2015. Specifically for the reimbursement code for Supartz/Supartz, it is projected that the Center for Medicare and Medicaid Services (CMS) will spend >$700 million over the next 5 years, and it is estimated that GenVisc 850 could save CMS $70–$100 million and patients >$8–$10 million in copayments over this same period through competitive pricing.

Similar to the experience with generic drugs, widespread availability of indistinguishable class III medical devices has the potential to reduce health-care costs while providing equivalent care. However, the first step in establishing such a benefit to the health-care system is that the FDA should establish scientifically valid and rigorous criteria for assessing such equivalence and embrace the legislation formally adopted in 1997 to allow such a process.7 The detailed clinical performance evaluations outlined in this paper and the approval of GenVisc 850 utilizing in part such analysis are the first attempt to validate the analytical methodology to establish such a standard for such evaluations and set the standard for future regulatory reviews.

Importantly, the Bayesian methodology used to support the GenVisc 850 PMA approval is complex and can be more difficult for clinicians and regulatory agencies to understand than the results of a classical frequentist statistical approach to single randomized clinical trials. However, it also has the advantage to consider all available data, as was the case in this report, and the decision to adopt this approach must incorporate a prospectively defined Statistical Analysis Plan. It should be noted that this plan contained more assumptions to be tested (eg, homogeneity over time and exchangeability) and to be verified than might be needed for a randomized clinical trial. The use of this methodology should not be considered the new standard for PMA approval or implemented without advance notice to and acceptance by regulatory agencies.

Conclusion

The preclinical and clinical evidence summarized herein supported FDA approval of a class III medical device based upon indistinguishable characteristics in comparison to an approved branded product. GenVisc 850 was determined to have: 1) noninferior efficacy compared with an FDA-approved HA, 2) superior efficacy vs saline injections, using novel Bayesian modeling, and 3) no unexpected safety risks. Increasing the availability of equivalent or indistinguishable class III medical devices will foster competition and innovation and at the same time significantly reduce health-care spending in the US.

Acknowledgment

The authors thank Doctor David Toledo-Velasquez, PhD, of OrthogenRx, Inc. and Doctor Pilar Colonel-Granado, PhD, for their facilitation in accessing the related data, comments, and contributions.

Disclosure

Doros, Lavin, and Miller were compensated by OrthogenRx, Inc. for their respective role in the conduct of this study. Daley is the President of OrthogenRx, Inc. The authors report no other conflicts of interest in this work.

References


