A comparative analysis of skin substitutes used in the management of diabetic foot ulcers

Objective: To compare the relative product cost and clinical outcomes of four skin substitutes used as adjunctive treatments for diabetic foot ulcers (DFUs).

Method: Medicare claims data from 2011 to 2014 were used to identify beneficiaries with diabetes and foot ulcers. Patients treated with one of four types of skin substitute (Apligraf, Dermagraft, OASIS, and MatriStem) were identified. The skin substitutes were compared on episode length; amputation rate; skin substitute utilisation; and skin substitute costs.

Results: There were 13,193 skin substitute treatment episodes: Apligraf (HML) was used in 4926 (37.3%), Dermagraft (HSL) in 5530 (41.9%), OASIS (SIS) in 2458 (18.6%) and MatriStem (UBM) in 279 (2.1%). The percentage of DFUs that healed at 90 days were: UBM 62%; SIS 63%; HML 58%; and HSL 58%. Over the entire time, UBM was non-inferior to SIS ($p<0.001$), and either was significantly better than HML or HSL ($p<0.005$ in all four tests). HML was marginally superior to HSL ($p=0.025$ unadjusted for multiple testing). Medicare reimbursements for skin substitutes per DFU episode for UBM ($1435 in skin substitutes per episode) and SIS ($1901) appeared to be equivalent to each other, although non-inferiority tests were not significant. Both were less than HML ($5364) or HSL ($14,424) ($p<0.0005$ in all four tests). HML was less costly than HSL ($p<0.0005$).

Conclusion: Various types of skin substitutes appear to be able to confer important benefits to both patients with DFUs and payers. Analysis of the four skin-substitute types resulted in a demonstration that UBM and SIS were associated with both shorter DFU episode lengths and lower payer reimbursements than HML and HSL, while HML was less costly than HSL but equivalent in healing.

Declaration of interest: MM and NM are health economic consultants who completed and authored this analysis. They have no financial interest in ACell, Inc. Their fees were not contingent on the outcome of this analysis.

cost-effectiveness analysis • diabetic foot ulcer • MatriStem • skin substitutes • wound healing

Among the nearly 30 million Americans with type 1 or type 2 diabetes mellitus (DM), the annual incidence of developing a diabetic foot ulcer (DFU) is estimated to be between 1–4%, with a lifetime risk of between 25–30%.[1,2,3] The overall direct medical costs of the treatment of each DFU may exceed $45,000.[4,5] In the US, the treatment of DFUs impose a substantial annual burden on public and private payers, ranging from $9–13 billion, which is in addition to the costs associated with treatments for the underlying diabetes.[6]

Healing and wound closure fail in 24–60% of DFUs.[7,8] DFUs are reported to account for up to two-thirds of all non-traumatic surgical lower limb amputations in the US.[3,9] In 2010, an estimated 73,000 lower limb amputations were performed in patients with DM, and ischaemic and infected DFUs were responsible for 25% of all hospital stays in patients with diabetes.[10]

Skin substitutes were developed to aid in the treatment and closure of wounds.[11] They are classified into allogeneic cell-containing skin substitutes (keratinocytes or fibroblasts), including autologous cell-containing skin substitutes, Apligraf (Organogenesis Inc., Canton, MA, US) and Dermagraft (Organogenesis Inc., Canton, MA, US), and acellular matrices including OASIS (Cook Biotech, West Lafayette, IN, US) and MatriStem (ACell, Columbia, MD, US).

A 2010 clinical consensus panel provided a series of recommendations which included the use of skin substitutes in conjunction with standard wound-care regimens.[12] The National Institute for Health and Care Excellence (NICE) 2015 guidelines recommended the use of skin substitutes as an adjunct to standard care when treating non-healing DFUs.[13] In 2016, the Cochrane Wounds Group published a systematic review on skin grafting and skin substitutes for the treatment of DFUs that included 1655 patients randomised in 17 clinical trials.[14] The authors concluded that, when compared with standard care alone, adjunctive treatment with skin grafts and skin substitutes could increase the healing rate of DFUs and lead to fewer amputations.[14]

Here, the results of a Medicare claims analysis of the costs and outcomes of DFUs treated with four types of skin substitutes are presented. The purpose of this analysis was to determine whether there are any reasons, in terms of effectiveness (quantity used per healing episode; healing/episode length; reduction in amputation outcomes) or costs for payers or providers, which would favour the use of one type of skin substitute over another.

M. Martinson,[1] MS, PhD, President; **N. Martinson**,[2] BS, President
E-mail: mmartinson@technomicsresearch.com
1 Technomics Research, LLC, Minneapolis, Minnesota, US. **2** Acclaim Data Analytics, LLC, Minneapolis, Minnesota, US.

Table 1. Sample selection criteria

Condition or definition	Inclusion criteria	Other criteria or comments
Diabetes	ICD-9 diagnosis codes: 250.00, 250.01, 250.40, 250.41, 250.50, 250.51, 250.60, 250.61, 250.70, 250.71, 250.80, 250.81, 250.90, 250.91	
Foot ulcer	ICD-9 diagnosis codes: 707.14, 707.15	
Index visit	A hospital inpatient or outpatient visit with the foot ulcer as the primary diagnosis code	A clean period of at least 60 days before the index visit, although it is possible that visits in the physician's office occurred before referral to the wound clinic or admission as an inpatient
Episode	• An episode length of at least 30 days from index visit to last visit; this was to eliminate easily healed wounds • At least one claim with a HCPCS code for one of four brands of skin substitutes (SS): Apligraf (Q4101), Dermagraft (Q4106), OASIS (Q4102, Q4103, Q4124), and MatriStem (Q4118, Q4119, Q4120). Q4118 is MicroMatrix, a micronised version of MatriStem. The weight of Q4118 used was converted into cm^2 of MatriStem for the analysis of units used. The conversion factor was 2.86mg/cm^2.	• A claim with a code for a brand not included in this list eliminated the episode, as did episodes with more than two of the listed brands • Hyperbaric oxygen therapy (HBOT) and negative pressure wound therapy (NPWT) were allowed during an episode as part of standard care as were debridement, bandages, and off-loading devices. Other advanced wound therapies during an episode, such as growth factors, eliminated the episode.
Foot amputation	**ICD-9 procedure codes:** 84.11 Amputation of toe 84.12 Amputation through foot	It is not possible in the claims to determine the side.
Leg amputation	**ICD-9 procedure codes:** 84.10 Lower limb amputation, not otherwise specified 84.13 Disarticulation of ankle 84.14 Amputation of ankle through malleoli of tibia and fibula 84.15 Other amputation below knee 84.16 Disarticulation of knee 84.17 Amputation above knee 84.18 Disarticulation of the hip 84.19 Abdominopelvic amputation	It is not possible in the claims to determine the side.

Methods

Data sources

The Centers for Medicare and Medicaid Services (CMS) Standard Analytical Files (SAF) from January 2011 to December 2014 were the source of the claims data. Patients in the dataset were de-identified, but hospitals were identifiable by name and location. Patient data in this database included demographic information on patient age and gender, medical diagnosis and procedures performed in the inpatient and outpatient hospital settings, emergency department and physicians' offices, and durable medical equipment used. Reimbursement for retail pharmacy drugs were not included in this analysis.

Since this dataset provided 100% of the hospital inpatient and outpatient visits, but only 5% of the physician office visits, a study limitation was to restrict the analysis to the claims of patients at hospitals with associated wound care clinics. The intention was to ensure that most of the wound care occurred in the clinic, and that most of the costs of treatment would be captured in the available claims. These hospitals were identified by first creating a list of the wound care clinics that treated 100 DFUs or more in 2012 and that had reported Current Procedure Terminology (CPT) codes for the application of skin substitutes. There were 814 hospitals that met these criteria. Patients who attended clinics at these hospitals between 2011 and 2014 were included.

Subject selection

The inclusion and exclusion criteria, including codes for subject selection, is included in Table 1, and is similar to the criteria used by Rice et al.[6] In brief, Medicare beneficiaries who had an inpatient or outpatient visit with foot ulcer as the primary diagnosis and a diagnosis of diabetes, but who had no prior foot ulcer claims for 60 days or more (i.e., a 60-day clean period) were included. Since office visits were not included in the dataset, it is possible that such visits occurred before referral to the wound clinic or admission as an inpatient. The hospital claim was flagged as the 'index event'. The episode had to have claims covering at least 30 days from the index hospital visit to eliminate easily healed wounds. The end of an episode was identified as a

wound-care claim followed by a clean period of 60 days. If an episode ended without an amputation during the episode, the DFU was considered healed.

The episode had to have at least one claim with a Healthcare Common Procedure Coding System (HCPCS) code identifying one of four types of skin substitutes chosen for this analysis: HML, HSL, SIS, or UBM. Skin substitutes are referred to by their type: Apligraf is a 2-layer skin substitute comprised of human keratinocytes and fibroblasts with a bovine Type 1 collagen lattice (HML); Dermagraft is a human fibroblast-derived dermal substitute with a bioabsorbable polyglactin mesh scaffold (HSL); the xenograft acellular matrices including OASIS, a porcine small intestine submucosal matrix (SIS); and MatriStem, a porcine urinary bladder matrix (UBM). The Q-codes from the SAF files are brand name-specific and do not refer to a technology or graft type for the product. In this analysis, abbreviations for the Q-code specific brand names will be used to reference the comparative group. A claim with a code for a brand not included in this list eliminated the episode from the analysis, as did episodes with more than two of the listed brands. This list was derived by evaluation of highest and lowest number of submitted claims for the types of skin substitutes.

Hyperbaric oxygen therapy (HBOT) and negative pressure wound therapy (NPWT) are commonly used therapies in the management of DFUs[15,16] and therefore were allowed during an episode. Other advanced wound therapies during an episode including, but not limited to, autologous platelet-rich plasma gel, stem cells, becaplermin gel, etc., eliminated the episode from the analysis as they are not yet considered standard of care (SOC) or widely used. They were used as exclusion criteria to determine the effect of skin substitutes independent of these adjunctive therapies.

Claims-based definitions

Similar to criteria used by Rice et al.,[6] the start of a skin substitute episode was the date of the first application of/treatment with a skin substitute; the end of an episode was date of the last claim before a minimum 60-day subsequent period in which there were no skin substitute claims for the treatment of a DFU. The episode length was the time between the beginning and end data of the application/treatment episode. If an episode ended with no amputation during that the treatment time or 60-day period, the DFU was considered to be healed. DFU episodes that were not healed by the end of the 2014 claims data were considered unresolved. Amputation was defined as an inpatient stay with an International Classification of Diseases 9th revision (ICD-9) procedure code for the amputation of a foot or leg. It is not possible to tell from the claims data that the DFU was the cause of the amputation, although this was the conservative assumption made for the analysis.

The number of skin substitute applications was defined as the number of days on which a claim with a HCPCS code for one of the four skin substitutes was present. The units of skin substitutes applied during the episode were calculated in cm^2. If claims for two of the four different types occurred sequentially, the treatment was considered to have been switched. Concomitant treatment occurred when two different skin substitute types commenced within one week of each other and overlapped for at least two weeks. If the time for initiating the second treatment was delayed by more than one week after the first treatment and overlapped for at least two weeks, the therapy was classified as an add-on.

Analysis

For continuous demographic variables, means were compared by the non-parametric analysis of variance, Kruskal–Wallis. For all categorical variables, the proportions were compared by Chi-squared (χ^2) tests.

For events occurring over time (wound healing, amputations), hazard ratios (HRs) were estimated from a Cox's proportional hazards regression analysis (Cox's regression) stratifying by whether one or two types of skin substitutes were used, whether osteomyelitis was present, and whether gangrene was present (since these were clinically and statistically significant). An HR is the ratio of the event rates in one group versus another.

Episode length was analysed using Cox's regression with yes/no covariates indicating which type of skin substitute was used first and stratified according to whether one or two types of skin substitutes were used. Death and amputations were censoring events. Foot amputations were analysed using Cox's regression stratified by whether one or two types were used, with death, healing or leg amputation as censoring events.

The standard statistical test comparing means, proportions, or HRs is an inequality test. Alternatively, if the comparison is intended to show that two treatments are the same, the appropriate test is a non-inferiority test. Such tests require a margin within which the two treatments would be considered clinically or financially equivalent. For costs, an additional $200 per episode was considered to be financially equivalent for the non-inferiority tests. For applications, two additional application days was considered to be equivalent; for units, an additional 25 cm^2 was considered to be equivalent. The non-inferiority margins for the HR for wound healing and amputations were set at 10% for the purposes of this analysis. Any comparative results within ±10% of each other were considered non-inferior.

All four types of skin substitutes were first compared to every other type using inequality tests (standard statistical tests of the null hypothesis of equivalence). If the inequality test was significant at α=0.05, one skin substitute of the pair was considered superior. If the inequality test was not significant (indicating that there was insufficient evidence that either was superior), then a non-inferiority test was conducted. This was a test of whether one treatment is within a margin of clinical (or financial) equivalence to the other. If this test was significant at α=0.05, the two skin substitutes of the pair were considered close enough to be equivalent.

Table 2. Diabetic foot ulcer episodes in the 2011–2014 Medicare data

Episodes	n (%) [95% CI]					p-value
	All skin substitutes	UBM	HML	HSL	SIS	
Episode count	13,193 (100%)	279 (2.1%) [1.9–2.4]	4926 (37.3%) [36.5–38.1]	5530 (41.9%) [41.1–42.7]	2458 (18.6%) [17.9–19.3]	χ^2 p=0.0003
Single-type	11,606 (100%)	232 (2.0%) [1.8–2.3]	4363 (37.6%) [36.7–38.5]	4898 (42.2%) [41.3–43.1]	2113 (18.2%) [17.5–18.9]	
Two-type	1587 (100%)	47 (3.0%) [2.2–4.0]	563 (35.5%) [33.1–37.9]	632 (39.8%) [37.4–42.3]	345 (21.7%) [19.7–23.8]	
Two-type episode count concomitant	NR* (100%)	NR* (NR)	NR* (NR)	33 (45.8%) [34.0–58.0]	35 (48.6%) [36.6–3.07]	—
Added-on: first type used	230 (100%)	11 (4.8%) [2.4–8.4]	79 (34.4%) [28.3–40.9]	86 (37.4%) [31.1–44.0]	54 (23.5%) [18.2–29.5]	χ^2 p=0.011
Added-on: second type used	230 (100%)	12 (5.2%) [2.7–8.9]	85 (37.0%) [30.7–43.6]	44 (19.1%) [14.2–24.8]	89 (38.7%) [32.4–45.3]	
Switched: first type used	1285 (100%)	36 (2.8%) [2.0–3.9]	480 (37.4%) [34.7–40.1]	513 (39.9%) [37.2–42.6]	256 (19.9%) [17.7–22.2]	χ^2 p<0.0001
Switched: second type used	1285 (100%)	52 (4.1%) [3.1–5.3]	403 (31.4%) [28.9–34.0]	322 (25.1%) [22.7–27.6]	508 (39.5%) [36.8–42.2]	

* Small numbers not reportable under the user agreement with CMS. UBM–MatriStem; HML–Apligraf; HSL–Dermagraft; SIS–OASIS; CI–confidence interval

The cost of each skin substitute application during the episode was estimated by multiplying the charges by the hospital-specific cost-to-charge ratio published by Medicare. All costs were inflated to 2014 US dollars using the medical care component of the Consumer Price Index (US Bureau of Labor Statistics).[17] This analysis was to compare the total skin substitute product costs from claims data and was not intended as a comprehensive evaluation of total cost of healing of DFUs.

The number of skin substitute applications, the amount of skin substitute used (cm^2), and the reimbursed amounts for skin substitutes (payer costs) were analysed using Wilcoxon's rank-sum test. If these tests were not significant, a non-inferiority test was constructed using a bootstrapped confidence interval of the differences, where bootstrapping is a resampling technique commonly used when data do not meet normality assumptions. The rationale for the testing is the same as for the clinical outcomes.

No data could be found which stated the consideration of equivalence in the characteristic by US payers. Given the magnitude of the mean values for each characteristic, the equivalence margin values for wound healing, average costs, application days or number of units seemed reasonable.

Results

Skin substitute treatment episodes

There were 13,193 skin substitute treatment episodes in the dataset, including 11,606 single-type episodes. In the 11,606 treatment episodes under analysis, HML was used in 4363 (37.6%); HSL was used in 4898 (42.2%); SIS in 2113 (18.2%), and UBM in 232 (2.0%). In the 1587 episodes in which two-types were used, HML was used in 563 (35.5%), HSL in 632 (39.8%), SIS in 345 (21.7%), and UBM in 47 (3.0%) indicating that HML or HSL were used more often in single-type episodes, while UBM and SIS were used more often in two-type episodes (p=0.0003).

Table 2 compares the four skin substitutes used in four application scenarios:

- Single type usage per episode
- Two types (of skin substitutes) used per episode
- Episodes in which treatments were switched within the time of treatment
- Episodes in which the second type was considered an add-on therapy. There were too few concomitant uses to allow a comparison to be made with a meaningfully powered statistical test.

Comparisons of patient and episode characteristics

The baseline characteristics of the patients studied and episode characteristics are shown in Table 3. Some of the comparisons were statistically significant but may not have been a difference that would be noticed in clinical practice, which can occur when the sample size is large enough to detect differences that may not be of any clinical consequence. This occurrence is not uncommon in claims analysis due to the large samples typically available. Note that age and days to first skin substitute application are examples of this. The percentage of patients reported within a given age category compared with the entire sample did not differ by more than 4%

Table 3. Patient and episode characteristics in the 2011–2014 data set

Characteristic	n (%) [95% CI] or mean (SD), [95% CI]					p-value
	All skin substitutes	UBM	HML	HSL	SIS	
Age						
<65	4485 (34.0%) [33.2–34.8]	88 (31.5%) [26.1–37.3]	1704 (34.6%) [33.3–35.9]	1921 (34.7%) [33.4–36.0]	772 (31.4%) [29.6–33.3]	
65–69	2788 (21.1%) [20.4–21.8]	68 (24.4%) [19.5–29.9]	1008 (20.5%) [19.4–21.7]	1219 (22.0%) [20.9–23.1]	493 (20.1%) [18.5–21.7]	
70–74	2012 (15.3%) [14.7–15.9]	38 (13.6%) [9.8–18.2]	730 (14.8%) [13.8–15.8]	847 (15.3%) [14.4–16.3]	397 (16.2%) [14.8–17.7]	χ^2 p=0.000
75–79	1613 (12.2%) [11.6–12.8]	29 (10.4%) [7.1–14.6]	586 (11.9%) [11.0–12.8]	700 (12.7%) [11.8–13.6]	298 (12.1%) [10.8–13.5]	
80–84	1231 (9.3%) [8.8–9.8]	27 (9.7%) [6.5–13.8]	480 (9.7%) [8.9–10.6]	471 (8.5%) [7.8–9.3]	253 (10.3%) [9.1–11.6]	
>84	1064 (8.1%) [7.6–8.6]	29 (10.4%) [7.1–14.6]	418 (8.5%) [7.7–9.3]	372 (6.7%) [6.1–7.4]	245 (10.0%) [8.8–11.3]	
Sex						
Female	4938 (37.4%) [36.6–38.2]	110 (39.4%) [33.6–45.4]	1865 (37.9%) [36.5–39.3]	2055 (37.2%) [35.9–38.5]	908 (36.9%) [35.0–38.8]	χ^2 p=0.732
Male	8255 (62.6%) [61.8–63.4]	169 (60.6%) [54.6–66.4]	3061 (62.1%) [60.7–63.5]	3475 (62.8%) [61.5–64.1]	1550 (63.1%) [61.2–65.0]	
Start of treatment						
Inpatient	704 (5.3%) [4.9–5.7]	13 (4.7%) [2.5–7.9]	286 (5.8%) [5.2–6.5]	292 (5.3%) [4.7–5.9]	113 (4.6%) [3.8–5.5]	χ^2 p=0.165
Outpatient	12,489 (94.7%) [94.3–95.1]	266 (95.3%) [92.1–97.5]	4640 (94.2%) [93.5–94.8]	5238 (94.7%) [94.1–95.3]	2345 (95.4%) [94.5–96.2]	
Days to use of skin substitute	63.9 (67.1) [65.1–62.8]	67.7 (88.4) [78.1–57.3]	65.2 (66.0) [67.0–63.4]	60.6 (63.7) [62.3–58.9]	68.2 (73.3) [71.1–65.3]	Kruskal–Wallis p=0.0001
Use of HBOT in episode?						
No	12,056 (91.4%) [90.9–91.9]	248 (88.9%) [84.6–92.3]	4552 (92.4%) [91.6–93.1]	5006 (90.5%) [89.7–91.3]	2250 (91.5%) [90.3–92.6]	χ^2 p=0.003
Yes	1137 (8.6%) [8.1–9.1]	31 (11.1%) [7.7–15.4]	374 (7.6%) [6.9–8.4]	524 (9.5%) [8.7–10.3]	208 (8.5%) [7.4–9.7]	
Use of NPWT in episode?						
No	12,600 (95.5%) [95.1–95.8]	262 (93.9%) [90.4–96.4]	4710 (95.6%) [95.0–96.2]	5266 (95.2%) [94.6–95.7]	2362 (96.1%) [95.3–96.8]	χ^2 p=0.188
Yes	593 (4.5%) [4.2–4.9]	17 (6.1%) [3.6–9.6]	216 (4.4%) [3.8–5.0]	264 (4.8%) [4.3–5.4]	96 (3.9%) [3.2–4.7]	
Cellulitis during episode?						
No	10,803 (81.9%) [81.2–82.6]	234 (83.9%) [79.1–88.0]	4076 (82.7%) [81.6–83.7]	4454 (80.5%) [79.4–81.5]	2039 (83.0%) [81.5–84.5]	χ^2 p=0.008
Yes	2390 (18.1%) [17.4–18.8]	45 (16.1%) [12.0–20.9]	850 (17.3%) [16.3–18.4]	1076 (19.5%) [18.5–20.6]	419 (17.1%) [15.6–18.6]	
Osteomyelitis during episode?						
No	10,201 (77.3%) [76.6–78.0]	196 (70.3%) [64.6–75.6]	3869 (78.5%) [77.3–79.6]	4223 (76.4%) [75.3–77.5]	1913 (77.8%) [76.1–79.4]	χ^2 p=0.002
Yes	2992 (22.7%) [22.0–23.4]	83 (29.8%) [24.5–35.5]	1057 (21.5%) [20.4–22.7]	1307 (23.6%) [22.5–24.7]	545 (22.2%) [20.6–23.9]	
Gangrene during episode?						
No	12,525 (94.9%) [94.5–95.3]	250 (89.6%) [85.4–92.9]	4687 (95.2%) [94.6–95.8]	5248 (94.9%) [94.3–95.5]	2340 (95.2%) [94.3–96.0]	χ^2 p=0.001
Yes	668 (5.1%) [4.7–5.5]	29 (10.4%) [7.1–14.6]	239 (4.9%) [4.3–5.5]	282 (5.1%) [4.5–5.7]	118 (4.8%) [4.0–5.7]	
Episode ended in amputation?						
No	11,975 (90.8%) [90.3–91.3]	250 (89.6%) [85.4–92.9]	4491 (91.2%) [90.4–92.0]	4969 (89.9%) [89.1–90.7]	2265 (92.2%) [91.1–93.2]	χ^2 p=0.006
Yes	1218 (9.2%) [8.7–9.7]	29 (10.4%) [7.1–14.6]	435 (8.8%) [8.0–9.6]	561 (10.1%) [9.3–10.9]	193 (7.9%) [6.9–9.0]	

UBM–MatriStem; HML–Apligraf; HSL–Dermagraft; SIS–OASIS; NPWT–negative pressure wound therapy; HBOT–hyperbaric oxygen therapy; CI–confidence interval

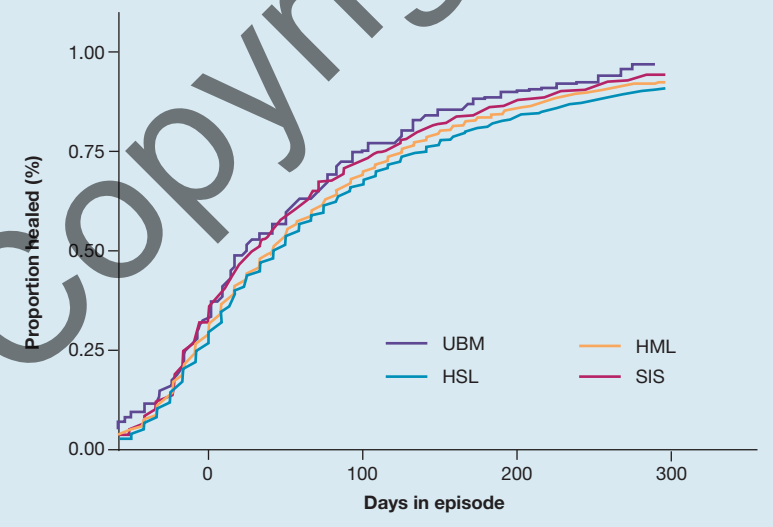
Table 4. Non-inferiority tests for time-to-healing and amputation

Event	Statistic	Comparator (denominator)	Event rate at 90 days. Hazard ratio numerator versus comparator [95% CI]			
			UBM	HML	HSL	SIS
Healing	%	None	62%	58%	58%	63%
	Cox's HR [95% CI], p-values: (1) inequality† (2) non-inferiority‡	UBM	NA	0.78 [0.68–0.88] (1) p<0.0005	0.74 [0.65–0.84] (1) p<0.0005	0.87 [0.76–0.99] (2) p<0.001
		HML		NA	0.95 [0.89–1.0] (1) p=0.025	1.1 [1.0–1.2] (1) p<0.005
		HSL			NA	1.2 [1.1–1.2] (1) p<0.0005
		SIS	1.1 [1.0–1.3] (2) p>0.2			NA
Amputation	%	None	2.1%	1.6%	1.4%	1.3%
	Cox's HR [95% CI], p-values: (1) inequality* (2) non-inferiority†	UBM	NA	0.83 [0.53–1.3] (2) p>0.2	0.98 [0.63–1.5] (2) p>0.2	0.88 [0.55–1.4 (2) p>0.2
		HML	1.2 [0.77–1.9] (2) p>0.1	NA	1.2 [1.1–1.3] (1) p>0.037	1.1 [0.93–1.2] (2) p=0.05
		HSL	1.0 [0.65–1.6] (2) p>0.2		NA	0.89 [0.74–1.1] (2) p>0.2
		SIS	1.1 [0.72–1.8] (2) p= 0.2	0.94 [0.76–1.2] (2) p>0.2	1.1 [0.91–1.4] (2) p=0.04	NA

* The hazard ratio is measured over the entire curve, not at 90 days when the percentage healed was reported. † Standard two-sided statistical test for the equality of the two event rates. Only one version of the inequality comparisons are shown because, for example, HML versus HSL is the same as HSL versus HML. ‡ Non-inferiority tests are one-sided by design. Using a relative 10% non-inferiority margin corresponding to a hazard ratio of 1.11 for healing because a hazard ratio over 1.0 means that the comparator has a lower healing rate, or 0.9 for amputation because a hazard ratio less than 1.0 means that the comparator has a higher amputation rate. Significant values mean that the comparator type has demonstrated non-inferiority (not superiority) to the other skin substitute. Two versions of the non-inferiority comparisons are shown because, for example, showing that HML is non-inferior to HSL is not the same test as showing that HSL is non-inferior to HML.
UBM–MatriStem; HML–Apligraf; HSL–Dermagraft; SIS–OASIS; NA–not applicable; CI–confidence interval

for any of the reports. The average time to starting skin substitute treatment after diagnosis of a DFU varied between a high of 68 days for UBM and a low of 61 days for HSL. This difference of seven days between starting the use of UBM and HSL is not a significant percentage of an elapsed time of more than two months (>60 days) until initiation of treatment with any skin substitute under analysis.

Fig 1. Time to diabetic foot ulcer healing



There were some clinical adverse events for which the bracketing difference in reported rates exceeded 5%, indicating a clinically significant difference. Osteomyelitis and gangrene were the most frequently reported adverse events. In all of the skin substitute treatment episodes with osteomyelitis or gangrene, the infection was present before any of the skin substitutes were applied.

Osteomyelitis was reported in 8.3% more episodes in which UBM (70.3%, 95% CI [64.6–75.6%]) was used than for HML treatment episodes (78.9%, 95% CI [77.3–79.6%]) (p=0.002), and gangrene was reported in 5.6% more episodes in which UBM treatment (89.6%, 95% CI [85.4–92.9%]) was used than for SIS treatment episodes (95.2%, 95% CI [94.3–96.0%]) (p=0.001).

Comparisons of outcomes

Fig 1 and Table 4 show the results of the comparative analysis of amputations and wound healing, including the HRs of the event rates. Table 4 shows the percentage of DFUs that healed at 90 days, the HRs over the entire follow-up time, and the p-values of the inequality

Table 5. Comparisons of cost and use. Reimbursements, applications and units used

	Statistic	Comparator skin substitute	Skin substitute			
			UBM	HML	HSL	SIS
Payer reimbursements	% Not reimbursed		38%	19%	8.9%	30%
	Mean (n, SD) [95% CI] if reimbursed		$1435 (174, $3160) [$965–1905]	$5364 (4010, $6966) [$5148–5580]	$14,424 (5040, $15074) [$14,008–14840]	$1901 (1721, $5394) [$1646–2156]
	Rank-sum p-value (1)* or Bootstrapped p-value (2)†	UBM	NA	(1) p<0.0005	(1) p<0.0005	(2) p=0.17
		HML		NA	(1) p<0.0005	(1) p<0.0005
		HSL			NA	(1) p<0.0005
		SIS	(2) p=0.67			NA
Number of applications	Mean (SD), [95% CI]		5.53 (7.88), [4.61–6.45][	3.24 (2.82), [3.16–3.32]	5.96 (4.64), [5.84–6.08]	4.48 (4.04), [4.32–4.64]
	Rank-sum p-value (1)* or Bootstrapped p-value (2)†	UBM	NA	(1) p=0.0001	(1) p=0.0001	(2) p<0.005
		HML		NA	(1) p<0.0005	(1) p<0.0005
		HSL			NA	(1) p<0.0005
		SIS	(2) p=0.023			NA
Units used	Mean (SD), [95% CI]		196 (494), [138–254]	190 (289), [182–198]	447 (516), [443–461]	155 (334), [142–168]
	Rank-sum p-value (1)* or Bootstrapped p-value (2)†	UBM	NA	(1) p=0.0001	(1) p=0.0001	(2) p=0.60
		HML		NA	(1) p<0.0005	(1) p<0.0005
		HSL			NA	(1) p<0.0005
		SIS	(2) p=0.005			NA

*Standard two-sided statistical test for the equality of the two costs, applications, or units. Only one version of the inequality comparisons are shown because, for example, HML versus HSL is the same as HSL versus HML. †Non-inferiority tests are one-sided by design. The non-inferiority margins were: a maximum cost difference of $200 per episode; a maximum number of application days of 2 per episode; a maximum number of units of 25cm^2 per episode. A significant value means that the comparator type has demonstrated non-inferiority (not superiority) to the other skin substitute. Two versions of the non-inferiority comparisons are shown because, for example, showing that HML is non-inferior to HSL is not the same test as showing that HSL is non-inferior to HML. The table shows that the data are sufficient to conclude that the number of units of UBM is no greater than 25cm^2 more than SIS (p=0.005), but there is insufficient evidence to conclude that the number of units of SIS is no greater than 25cm^2 more than UBM (p=0.60).
UBM–MatriStem; HML–Apligraf; HSL–Dermagraft; SIS–OASIS; NA–not applicable; SD–standard deviation; CI–confidence interval

(labelled '(1)') and non-inferiority (labelled '(2)') tests. HRs less than 1.0 mean a higher rate of healing or amputation with the comparator brand.

Overall, healing rates were significantly lower for HML and HSL compared with UBM (HRs 0.78 and 0.74) and SIS (HRs 0.91 and 0.83) (p<0.005 on all four tests). HSL had significantly lower rates than HML (HR 0.95, p=0.025). SIS and UBM were not significantly different in the inequality test, but UBM was non-inferior to SIS (HR 0.87, p<0.001). Healing at 90 days was 62% for UBM, 63% for SIS, and 58% for HML and HSL.

In this analysis, amputation was a censoring event, meaning that healing could not be observed if the foot or leg was amputated. Because it would be possible for a skin substitute with a high rate of amputation to appear to be superior or equivalent in healing due to censoring, it was important to test the amputation rates. Table 4 shows that the adjusted amputation rates were all within a range of approximately 0.8%, ranging from 2.1% at 90 days with UBM, to 1.3% with SIS.

Of all of the amputation tests, only the p-values for the inequality test of HSL to HML (HR 1.2, p=0.04) and

the non-inferiority test of HSL to SIS (HR 1.1, p=0.04) came close to significance. However, because there are a large number of tests, most analysts would consider it important to adjust the p-value used for significance for multiple testing (also called 'multiple inference'). The reason is that every significance test has a chance of being wrong, and the compound probability of being wrong on at least one of many tests can become quite high. If the p-value of significance is adjusted for the fact that there are 18 tests in Table 4 (a typical adjustment would be to 0.05/18=0.003), these two tests would not be significant.

Comparison of skin substitute applications and amounts

The number of applications of skin substitute was compared among the four types. The mean number of applications of each type was: 5.5±7.9 applications for UBM; 6.0±4.6 applications of HSL; 3.2±2.8 applications for HML (Table 5); and 4.5±4.0 for SIS. The mean number of UBM applications was significantly more than HML (p<0.0001); but there were significantly fewer applications of UBM than HSL (p<0.0001). SIS also had significantly fewer applications than either HML or HSL (p<0.0005 for both tests). HML had fewer than HSL (p=0.0001), while SIS was non-inferior compared with UBM (p<0.005) and there was some indication that UBM was also non-inferior to SIS (p=0.023) (Table 5).

The total number of units of skin substitute (cm^2) applied during a treatment episode was compared: UBM had 196±494 units applied; HML had 190±289 units applied; HSL had 447±516 units applied; and SIS had 155±334 units applied on average during a treatment episode. The applied units of UBM were significantly more than HML (p=0.0001) but significantly less than HSL (p=0.0001). SIS units were significantly less than HML and HSL (p<0.0005 in both tests) and non-inferior to UBM (p= 0.005). HML units applied were significantly less than HSL units (p<0.0005).

Comparison of skin substitute costs

In the US, payer coverage for skin substitutes can vary. If a skin substitute is listed by a payer as covered, then payment or reimbursement for the skin substitute is possible. HSL, HML, and SIS are commonly covered and reimbursed by Medicare for lower extremity wounds such as DFUs, whereas UBM is covered in some geographic jurisdictions and not others. In reviewing the data, 8.9% of HSL episodes included no reimbursement for skin substitutes, while 19% of HML, 30% of SIS and 38% of UBM episodes had no reimbursement for skin substitutes. Because of the high percentage of non-reimbursements across these four types of skin substitutes, average reimbursements were compared only among episodes with skin substitute reimbursements.

The mean (± standard deviation) of the reimbursed amounts were as follows: UBM: $1435±3160; HML: $5364±6966; HSL: $14,424±15,074; and SIS: $1901±5394 (Table 5). Both UBM and SIS were significantly less expensive than HML and HSL (p<0.0005 in all four tests). HML was significantly less expensive than HSL (p<0.0005). Neither UBM nor SIS were significant in the non-inferiority tests with the other.

Discussion

DFUs that do not heal with standard care alone can be successfully treated with skin substitutes. However, these skin substitutes can be expensive, and once their use is initiated, treatment may still be required for several months. In total, these episodes can be long and expensive, so therapies that can decrease the duration of the episode without creating a financial burden should be encouraged.

A recently published systematic review from the Cochrane Wounds Group concluded that the clinical effectiveness and cost-effectiveness of skin substitutes in the treatment of DFUs still remains uncertain and that no specific type of skin substitute could be recommended.[18] The present study has begun to address this uncertainty.

This study has provided data on clinically-based evidence from CMS claims that two skin substitutes, HSL and HML, did not appear superior to the lesser-used types of skin substitute, UBM and SIS. At 90 days, UBM and SIS healed about 62–63% of DFUs, while HML and HSL healed about 58%, which was statistically significant due to the large sample sizes available in the claims data.

The 4% difference in healed wounds may not be considered significant to the clinicians using skin substitutes but the cost difference could be compelling to clinicians and payers. While HML and HSL cost over $5000 to heal a DFU, UBM and SIS were less than $2000. If payers can save even a portion of that difference on every DFU on which skin substitutes are used, the financial benefit will be substantial.

Limitations

There were several limitations to this study. While the patients were compared on many baseline characteristics and were found to be mostly equivalent, there may have been some systematic differences that were not measured or discernible in the SAF data sets. The two characteristics that were both statistically and clinically significantly different, osteomyelitis and gangrene, were controlled in the analysis.

Episode costs were limited to reimbursements for skin substitutes applied in a hospital setting. This was necessary because it was not possible to estimate the cost of the entire episode of care. The cost of inpatient and outpatient drug treatment during these episodes was excluded from all analysis, as the emphasis was on skin substitute use costs and not total cost of healing. The Medicare data does not include pharmacy claims and only limited office-visit claims. Other costs that were not included were costs borne by

supplemental insurers and out-of-pocket costs paid by the patient. Further studies to analyse total episode costs of direct medical care should be undertaken.

Conclusion

This is the first US Medicare claims study to compare clinical outcomes and costs of four types of skin substitutes in the treatment of DFUs. The two types of skin substitute predominately used, HML and HSL, did not perform as well as the lesser-used types, UBM and SIS, in terms of DFU treatment episode length, as approximately 4% fewer DFU episodes were healed at 90 days. In addition, the two leading skin substitutes were more expensive by a minimum of $3000 per episode. **JWC**

Acknowledgements:
Dr Dinah V. Parums, Freelance Medical & Scientific Editor & Writer

References

1 American Diabetes Association. Clinical Practice Recommendations: foot care in patients with diabetes mellitus. Diabetes Care 1996; 19: S23.
2 Singh, N., Armstrong, D.G., Lipsky, B.A. Preventing foot ulcers in patients with diabetes. JAMA 2005; 293: 217–228.
3 Gregg, E.W., Gu, Q., Williams, D. et al. Prevalence of lower extremity diseases associated with normal glucose levels, impaired fasting glucose, and diabetes among U.S. adults aged 40 or older. Diabetes Res Clin Pract 2007; 77: 3, 485–488.
4 Jeffcoate, W.J., Harding, K.G. Diabetic foot ulcers. Lancet 2003; 361: 9368, 1545–1551.
5 Stockl, K., Vanderplas, A., Tafesse, E., Chang, E. Costs of lower-extremity ulcers among patients with diabetes. Diabetes Care 2004; 27: 2129–2134.
6 Rice, J.B., Desai, U., Cummings, A.K. et al. Burden of diabetic foot ulcers for Medicare and private insurers. Diabetes Care 2014; 37: 3, 651–658.
7 Hinchliffe, R.J., Valk, G.D., Apelqvist, J. et al. Specific guidelines on wound and wound-bed management. Diabetes Metab Res Rev 2008; 24: S1, S188–S189.
8 Game, F.L., Hinchliffe, R.J., Apelqvist, J. et al. A systematic review of interventions to enhance the healing of chronic ulcers of the foot in diabetes. Diabetes Metab Res Rev 2012; 28: S1, 119–141.
9 Ramsey, S.D., Newton, K., Blough, D. et al. Incidence, outcomes, and cost of foot ulcers in patients with diabetes. Diabetes Care 1999; 22: 382–387.
10 American Diabetes Association Statistics about diabetes. 2014. www.diabetes.org/diabetes-basics/statistics/ (accessed 19 September 2016).
11 Marston, W.A., Hanft, J., Norwood, P. et al. The efficacy and safety of Dermagraft in improving the healing of chronic diabetic foot ulcers: results of a prospective randomized trial. Diabetes Care 2003; 26: 1701–1705.
12 Snyder, R.J., Kirsner, R.S., Warriner, R.A. et al. Consensus recommendations on advancing the standard of care for treating neuropathic foot ulcers in patients with diabetes. Ostomy Wound Manage 2010; 56: S1–S24.
13 National Institute for Health and Care Excellence. Diabetic foot problems: prevention and management. NICE guideline NG19. NICE, 2015. http://tinyurl.com/hnjuqrt (accessed 19 September 2016).
14 Santema, T.B., Poyck, P.P., Ubbink, D.T. Skin grafting and tissue replacement for treating foot ulcers in people with diabetes. Cochrane Database Syst Rev 2016; 2, CD011255.
15 Sing, S., Gambert, S.R. Hyperbaric oxygen therapy: a brief history and review of its benefits and indications for the older adult patient. Annals of Long-Term Care: Clinical Care and Aging 2014; 22: 7/8, 37–42. http://tinyurl.com/jjgvkp9 (accessed 19 September 2016).
16 Hasan, M.Y., Teo, R., Nather, A. Negative-pressure wound therapy for management of diabetic foot wounds: a review of the mechanism of action, clinical applications, and recent developments. Diabet Foot Ankle. 2015; 6: 27618. doi: 10.3402/dfa.v6.27618.
17 U.S. Bureau of Labor Statistics. Annual Consumer Price Index website: www.bls.gov/cpi/home.htm (accessed March 2016).
18 Santema, T.B., Poyck, P.P.C., Ubbink, D.T. Editorial Group: Cochrane Wounds Group (2016). Skin grafting and tissue replacement for treating foot ulcers in people with diabetes. www.cochrane.org/CD011255/WOUNDS_skin-grafting-and-tissue-replacement-treating-foot-ulcers-people-diabetes (accessed March 2016).