Use of flexibility tests in the manufacturing process of 60° Björk-Shiley convexo-concave valves and the risk of outlet strut fracture

Rumana Z. Omar, PhD^{a,b,c}
 Linda S. Morton, RN^a
 Shahed Murad, MSc^d
 Kenneth M. Taylor, FRCS^a

Objectives: Outlet strut fracture remains a concern for 30,000 patients living with a Björk-Shiley convexo-concave heart valve (Shiley, Inc, Irvine, Calif, a subsidiary of Pfizer, Inc). Previous studies (Netherlands and United Kingdom) investigating valve manufacturing aspects identified multiple performance of the hook deflection test as a risk factor for 60° valves. The present study validated this finding using new data with a greater number of valves implanted worldwide. Risks of outlet strut fracture associated with other manufacturing aspects were also investigated.

Methods: A matched case-control study design was used including 416 outlet strut fracture cases and 803 controls.

Results: Analyses similar to that of the Dutch and United Kingdom studies produced odds ratios of 3.4 (95% confidence interval [CI]: 1.1-10.3) and 2.8 (95% CI: 1.1-7.3), respectively, for multiple hook deflection tests. Load deflection test, which replaced the hook deflection test, showed a statistically significant association with outlet strut fracture: odds ratio of 5.0 (95% CI: 2.1-11.8) and 6.2 (95% CI: 2.2-18.0) for single and multiple load deflection tests, respectively. An analysis where hook deflection tests were separated from load deflection tests showed significantly elevated odds ratios with performance of any type of flexibility test, and the highest odds ratio was observed with a combined performance of load and hook deflection tests.

Conclusions: Multiple hook deflection tests can now be considered for inclusion in the risk model used for guidelines on explant surgery to improve prediction of outlet strut fracture and provide patient reassurance. Load deflection tests and combined performance of hook and load deflection tests were found to be significant risk factors. No outlet strut fractures were reported for valves manufactured after March 1984 when the load deflection test was still in place. Examining manufacturing documents for these valves may identify new risk factors that could be responsible for the outlet strut fractures risk that remains unexplained to date.

From the Department of Cardiac Surgery, Imperial College School of Medicine, Hammersmith Hospital,^a the Department of Statistical Science, University College London,^b the Medical Statistics Unit, R & D Directorate, University College Hospital NHS Trust,^c and the Department of HIV and GU Medicine, Kings College London,^d London, United Kingdom.

Supported by the Trustees for the Bowling Pfizer Heart Valve Settlement Funds.

Received for publication Oct 10, 2002; revisions requested Nov 18, 2002; revisions received Feb 5, 2003; accepted for publication Feb 11, 2003.

Address for reprints: Professor Kenneth M. Taylor, Hammersmith Hospital, Imperial College School of Medicine, Department of Cardiac Surgery, Du Cane Road, London W12 0NN, United Kingdom (E-mail: k.m.taylor@ic.ac.uk).

J Thorac Cardiovasc Surg 2003;126:832-6

Copyright © 2003 by The American Association for Thoracic Surgery

0022-5223/2003 \$30.00 + 0

doi:10.1016/S0022-5223(03)00362-3

There are currently approximately 30,000 living patients worldwide with implanted Björk-Shiley convexo-concave (BSCC) heart valves (Shiley, Inc, Irvine, Calif, a subsidiary of Pfizer, Inc). Fracture of the valve's outlet strut, usually resulting in the patient's death, has been widely reported in medical journals and the international press since the 1990s. Development of potential diagnostic tools has been explored and epidemiologic research has been conducted to identify patient and valve characteristics that could be associated with the risk of outlet strut fracture (OSF).¹⁻⁶ This has enabled the development of guidelines for clinicians and BSCC patients for decisions on valve explant surgery or for patient reassurance.

Studies conducting detailed investigation on various aspects of the manufacturing process for 60° BSCC valves implanted in the United Kingdom and The Netherlands^{7,8} have identified multiple performance of an outlet strut flexibility test known as the hook deflection test (HDT) as a risk factor. The HDT was introduced to test for deficient welds. In this test, the valves were placed in a testing apparatus, a 5-kg weight was suspended from the outlet strut for less than 21-mm valves and a 7-kg weight for more than 21-mm valves, and the deflection of the outlet strut measured. Subsequently, the test was allowed to be performed more than once after rework (usually reweld) done on the valve. HDT was later replaced with the load deflection test (LDT), in which valves were placed in a specialized testing apparatus and a 5-kg load gradually was placed on the outlet strut, the weight was unloaded, and then it was reapplied. Recalled valves and valves manufactured during the period of changeover from HDT to LDT may have undergone both tests. The time periods for flexibility tests in the manufacturing process are summarized below.

1. Before May 1980: no flexibility test was performed.
2. May 1980 to August 1981: HDT was performed only once.
3. August 1981 to February 1982: one or more HDT was performed.
4. February 1982 to April 1984: one or more LDT was performed, HDT and LDT were performed if the valve was recalled or manufactured during the period of changeover.
5. April 1984: LDT performance continued as well as new checking procedures for disc-to-strut gap (DSG).

The previous findings of performance of multiple HDT as a risk factor were based on a small number of valves (6 multiple HDT in the United Kingdom with 4 OSFs, 14 in The Netherlands with 4 OSFs), and the confidence intervals for estimates of association were very wide. Furthermore, some of the multiple HDT valves that were subjected to an OSF had also undergone an LDT. Although in the United Kingdom study⁸ the association between OSF and performance of LDT was investigated, it was not possible to disentangle the effects of the two types of flexibility tests because of small numbers. LDT was not investigated in the Dutch study. The present study was conducted to validate the finding of performance of multiple HDTs as a risk factor, using new data with a greater number of OSF valves, and to improve the precision of previously reported associations for multiple HDTs. We also examined the association with HDT and LDT separately. A secondary objective was to investigate other aspects of the manufacturing process that were observed as risk factors for OSF in previous manufacturing studies.^{7,8}

Methods

The characteristics of the valve manufacturing process have been described in depth elsewhere.⁸ Risk of OSF was investigated with respect to the following aspects of the manufacturing process:

1. Use of HDT and LDT flexibility tests
2. Presence of anomalies in the valve, such as crack
3. Maximum clearance from disc to outlet strut at any opening angle
4. Presence of a phantom welder
5. Extent of rework done on the valve such as reweld and number of outlet strut manipulations

Cases and controls were identified from the manufacturers research database. Sixty-degree valves not included in the previous manufacturing studies,^{7,8} with known geographic location and documented proof of implant, were eligible for inclusion. Valves with erroneous implant dates were excluded. Using an end date of December 31, 2000, a total of 416 60° valves implanted worldwide that had experienced an OSF were included as cases. Two controls were selected for each case matched on valve size, position, and geographic location of implant. For 15 cases only 1 control could be found. Controls could not be found for 1 case. Information on factors described above (in points 1-5) was obtained from the valves' manufacturing records. There were 42 valves with some manufacturing documents missing and 2 with all documents missing. Data were abstracted, by 2 independent abstractors, in an identical manner to the previous studies.^{7,8} Information on age at implant, shoporder fracture rate, manufacturing date, and welder groups were obtained from the manufacturer's research database.

Statistical Analysis

A conditional logistic regression analysis was performed to examine the association between potential risk factors and OSF. Information on age at implant was missing for 17% of the cases and 27% of the controls. Multiple imputation⁹ and mean imputation techniques were used to substitute missing values of age. Multiple regression models matching closely to those reported in previous manufacturing studies were used to validate the finding of multiple HDT as a risk factor. Retaining previously identified risk factors (patients' age at implant, percentage shoporder fracture rate, manufacturing date, and welder group) in the regression model, a backward elimination procedure was used to examine the remaining manufacturing factors.¹⁰ Regression models also were examined including a factor defined as flexibility test, where LDT and HDT were separated and categorized as follows: (1) no HDT or LDT; (2) HDT only; (3) >1 HDT; (4) LDT only; (5) >1 LDT; (6) any combination of HDT and LDT (1 HDT + 1 LDT, >1 HDT + 1 LDT, >1 LDT + 1 HDT, >1 HDT + >1 LDT).

Valves were selected from a nonmonitored population and cases and controls could not be matched on duration of implant. A sensitivity analysis was conducted by applying the same model on a subset of cases and controls, closely matched on time of implant. Resampling methods were used to check the stability of models based on the backward elimination method.¹¹ All analyses were performed using the statistical software STATA 7.0 (Stata Corp, College Station, Tex).

TABLE 1. Distribution of valves with type of flexibility tests

Flexibility tests	Cases/controls	Total
No test	34/210	244
1 HDT only	153/420	573
>1 HDT only	16/14	30
1 LDT only	123/100	223
>1 LDT only	31/18	49
Combined HDT and LDT	57/42	99

Information on flexibility tests missing for 3 valves (2 cases and 1 control).

Results

The number of valves undergoing the 2 types of flexibility tests is presented in Table 1. There were 244 valves manufactured before May 1980 that had not undergone any flexibility test. Of the 99 valves that had undergone both HDT and LDT, 31 were recall valves. Multiple regression models matching as closely as possible to those reported in the Dutch (model 1) and United Kingdom (models 2 and 3) studies are presented in Table 2. The results show that HDT is significantly associated with the risk of OSF, as found in the Dutch and United Kingdom studies.^{7,8}

In a multiple regression model including all the manufacturing factors investigated in this study except HDT, only LDT remained statistically significant after accounting for previously established risk factors (Table 3). Both single and multiple performances of LDT showed a statistically significant association with OSF.

Results from a model where HDT and LDT were treated as separate categories of the flexibility test are shown in Table 4. The baseline comparison group consists of valves that did not undergo any flexibility test. The results indicate that the performance of any type of flexibility test is associated with OSF, with the strongest association observed for performance of a combination of LDT and HDT. The confidence intervals associated with the various categories of flexibility test are wide, as the numbers are small due to partitioning into finer categories.

LDT and flexibility test were found to be strong predictors of OSF when checking model stability (chosen in 70% and 76% of the repeated samples, respectively).

Similar results were obtained when missing age was replaced with mean age. Flexibility test remained statistically significant in the sensitivity analysis performed on a subset of valves closely matched on implant date with estimated odds ratios somewhat smaller than that observed for the complete set. However, the power of the study was substantially reduced for this analysis.

Discussion

Over the past decade, research has been continued for BSCC valve patients to identify patients at a high risk for valve fracture. The primary objective of this study was to validate the performance of multiple HDT as a new risk

factor. The current study uses data with the largest number of OSFs reported to date among 60° valves and includes valves implanted worldwide. The odds ratio for multiple HDTs was 2.8 (95% CI: 1.1-6.7) and may be compared with an odds ratio of 7.0 (95% CI: 1.1-45.5) observed in the United Kingdom.⁸ The odds ratio for multiple HDTs for large-size mitral valves was 3.4 (95% CI: 1.1-10.3), which may be compared with an odds ratio of 6.2 (95% CI: 1.9-21) observed in the Dutch study.⁷ The confidence intervals in the previous studies overlap with each other and also with that of the current study. Multiple HDT as a risk factor has been confirmed in the present study and the estimates of odds ratios have greater precision but are lower than those observed previously.

Valves that failed the first HDT were usually recorded to have an anomaly on or near the outlet strut such as a crack. The presence of anomalies did not show any statistically significant association with OSF in the multiple regression analysis. Performance of multiple HDT can now be considered for inclusion as a risk factor in the model currently used to produce guidelines for elective explant surgery. With an aging BSCC valve population, the risks of mortality and morbidity associated with elective explantation and re-replacement of valves are likely to increase. Improved prediction of OSF may lead to the identification of high risk groups of patients so that appropriate advice could be provided to them on explant surgery and reassurance given to patients with valves that have relatively lower risks of OSF.

Multiple HDTs were in place during August 1981 to February 1982 and OSFs have been reported for valves manufactured outside this time frame. LDT, which replaced HDT, has been identified as a risk factor for OSF with an odds ratio of 5.0 (95% CI: 2.1-11.8) and 6.2 (95% CI: 2.2-18.0) for single and multiple performances, respectively. An increased risk with LDT was observed in the United Kingdom study, but the estimates of odds ratios were smaller (OR: 2.8, 95% CI: 1.5-6.8 for single LDT; OR: 1.2, 95% CI: 0.3-6.4 for multiple LDT, from univariate analysis). The current study had a sufficient number of valves to enable an investigation of the risks associated with LDT and HDT separately and also in combination. The results from this analysis suggest that all flexibility tests are associated with the risk of OSF, in particular the performance of HDT combined with LDT. A relatively low risk observed with a single HDT in previous studies could be due to the fact that the baseline group used for comparison included those valves that had undergone an LDT, a significant risk factor, resulting in an underestimation of the associated risk.

LDT appears to be a more complex quality assurance test involving interpretation of graphs. From the engineering specifications, it appears that the test could result in either the valve failing, requiring rework and then being retested, or the graph results indicating that the valve was incorrectly

TABLE 2. Results from multiple regression analysis investigating the risk of fracture with HDT

Model	Not adjusted for age			Adjusted for age*	
	Risk factor	OR	95% CI	OR	95% CI
Model 1 (Dutch model, based on ≥ 29 mm mitral valves)	> 1 HDT†	3.5	1.4–8.7	3.4	1.1–10.3
	More than one disc fitted‡	1.0	0.7–1.6	1.0	0.6–1.7
	Shoporder fracture rate (%)	1.4	1.3–1.5	1.4	1.3–1.5
	Age at implant (years)	N/A	N/A	0.96	0.94–0.97
Model 2 (UK model)	1 HDT§	0.7	0.5–1.0	0.8	0.5–1.2
	> 1 HDT	2.4	1.1–5.1	2.8	1.1–7.3
	Disc-to-strut gap after placement of sewing ring (inches)				
	0.012	1.5	1.1–2.0	1.4	1.0–2.1
	0.013	1.5	1.0–2.2	1.5	0.9–2.4
	> 0.013	1.9	1.2–2.9	1.9	1.2–3.4
	Shoporder fracture rate (%)	1.3	1.2–1.4	1.3	1.2–1.4
	Age at implant (years)	N/A	N/A	0.95	0.94–0.97
	1 HDT§	0.7	0.5–1.0	0.8	0.5–1.1
	> 1 HDT	2.4	1.1–5.0	2.5	1.0–6.8
Model 3 (UK model)	Disc-to-strut gap before placement of sewing ring (inch)				
	0.012	1.2	0.9–1.7	1.1	0.7–1.7
	0.013	1.5	1.0–2.4	1.3	0.8–2.3
	> 0.013	1.1	0.7–1.6	1.1	0.7–1.9
	Shoporder fracture rate (%)	1.3	1.2–1.4	1.3	1.2–1.4
	Age at implant (years)	N/A	N/A	0.95	0.94–0.97

*Replacing missing age with age imputed from multiple imputation.

†Baseline: no HDT (includes valves with no flexibility tests + LDT only) + 1 HDT.

‡Baseline: no discs fitted.

§Baseline no HDT (includes valves with no flexibility test + LDT only).

||Baseline: <0.012 inch.

placed in the testing apparatus. The number of tests is recorded if the valves had undergone multiple tests for test failure. If the valve has been incorrectly fitted into the apparatus and underwent retesting, only one LDT is recorded. This could account for the association observed with the performance of one LDT. If a valve failed the test it is recorded as reworked with no record of the type of rework performed.

Information was available on the valve implant dates for both cases and controls and the OSF dates for cases; however, mortality or valve explant dates were not known for controls. Therefore, the cases and controls could not be matched on duration of implant. This could introduce some bias in the estimates of the odds ratios for the flexibility test as it is time dependent. However, the flexibility tests remained statistically significant in the subset analysis of valves closely matched on implant date. Since controls were selected from nonmonitored populations, there is a possibility of underreporting of OSFs, which could be higher among aortic valves.^{5,12,13} Although patient age, which is an important predictor for OSF, was not available for all valves, similar results were obtained when using different methods of age imputation and an analysis excluding age.

TABLE 3. Results from multiple regression analysis investigating the risk of fracture with LDT

Risk factor	Not adjusted for age		Adjusted for age*	
	OR	95% CI	OR	95% CI
1 LDT†	6.1	3.0–12.4	5.0	2.1–11.8
> 1 LDT	7.3	3.1–17.2	6.2	2.2–18.0
Manufacturing date‡				
1979	1.4	0.4–4.5	1.0	0.2–4.9
1980	1.8	0.5–6.3	1.1	0.2–5.2
1981	17.5	4.8–63.5	10.9	2.3–51.2
1982 to March 1984	28.6	7.4–110.8	22.5	4.4–114.6
Shoporder fracture rate (%)	1.2	1.1–1.2	1.2	1.1–1.3
Welder group C§	1.7	1.1–2.6	1.6	0.9–2.6
Age at implant (years)	N/A	N/A	0.95	0.94–0.97

*Replacing missing age with imputed age from multiple imputation.

†Baseline: no LDT (includes no flexibility test + HDT only).

‡Baseline: before 1979.

§Welder groups A and B (welders were divided into groups A, B, and C according to the risk of OSF associated with each welder).⁶

Moreover, the estimates associated with well-established risk factors such as age and shoporder fracture rate in this study were similar to those observed previously.^{5–7} Thus,

TABLE 4. Results from multiple regression analysis investigating the risk of OSF with flexibility tests separated into specific categories

Risk factor	Not adjusted for age		Adjusted for age*	
	OR	95% CI	OR	95% CI
Flexibility tests†				
1 HDT	3.5	1.3–9.7	3.8	1.2–12.0
> 1 HDT	9.6	2.3–40.5	9.6	1.9–60.7
1 LDT	20.5	5.5–76.9	16.8	3.6–85.6
> 1 LDT	20.3	4.8–86.4	17.7	3.1–106.8
Combined LDT and HDT	24.1	7.1–82.4	20.3	4.9–87.8
Manufacturing date‡				
1979	1.5	0.4–5.3	1.2	0.2–6.0
1980	1.1	0.3–4.3	0.6	0.1–3.5
1981	8.3	2.0–34.5	4.7	0.9–25.8
1982 to March 1984	16.4	3.7–71.4	12.1	2.1–68.7
Shoporder fracture rate (%)	1.15	1.06–1.25	1.17	1.05–1.30
Welder group C§	1.6	1.1–2.4	1.5	0.9–2.6
Age at implant (years)	N/A	N/A	0.95	0.94–0.97

*Replacing missing age with imputed age from multiple imputation.

†Baseline: no flexibility test.

‡Before 1979.

§Welder groups A and B.

the issue of missing age did not affect the conclusions from this study. Similar analyses were also performed for 70° valves and none of the manufacturing risk factors showed any statistically significant association with OSF. The results were not discussed any further because the main objective of this study was to validate previous findings that were based on 60° valves.

All valves produced from May 1980 onwards would have undergone flexibility tests, which is the vast majority of all BSCC valves (estimated to be 86% from the manufacturer's database used to select the case-control sample in this study). However, no OSFs were reported for valves manufactured after March 1984 (comprising 9.5% of valves) when LDT was still in place but disc-to-strut clearances underwent new checking procedures. These valves would, therefore, provide a pure comparison group to gain more insight into unexplained risks of OSFs. It was not possible to investigate the relationship between LDT and

disc-to-strut gap measurements in this study as there are no abstracted data for valves produced after March 1984. Investigation of manufacturing documents for these valves may help in identifying additional contributory risk factors and aid in further refinement of the guidelines for explant surgery.

We thank Dr. Gareth Ambler and Dr. Julie Barber for many helpful discussions. We thank Victoria Morton and Fizzah Chowdhury for help with data abstraction and entry.

References

1. Lindblom D, Rodriguez L, Bjork VO. Mechanical failure of the Björk-Shiley valve. *J Thorac Cardiovasc Surg.* 1986;92:894-907.
2. Van der Graaf Y, De Waard F, Van Herwerden LA, Defauw J. Risk of strut fracture of Björk-Shiley valves. *Lancet.* 1992;339:257-61.
3. Walker AM, Funch DP, Sulsky SI, Dreyer NA. Patient factors associated with strut fracture in Björk-Shiley 60° convexo-concave heart valves. *Circulation.* 1995;92:3235-9.
4. Walker AM, Funch DP, Sulsky SI, Dreyer NA. Manufacturing characteristics associated with strut fracture in Björk-Shiley 60° convexo-concave heart valves. *J Heart Valve Dis.* 1995;4:640-8.
5. Omar RZ, Morton L, Halliday D, Darns E, Beirne M, Blot WJ, et al. Outlet strut fracture of Björk-Shiley convexo-concave heart valves: the UK cohort study. *Heart.* 2001;86:57-62.
6. Blot WJ, Omar RZ, Kallewaard M, Morton LS, Fryzek, JP, Ibrahim MA, et al. Risk of fracture of Björk-Shiley 60° CC prosthetic heart valves: long-term cohort follow up in the Netherlands, the United Kingdom, and United States. *J Heart Valve Dis.* 2001;10:202-9.
7. Kallewaard M, Algra A, Defauw J, Van der Graaf Y. Which manufacturing characteristics are predictors of outlet strut fracture in large 60° Björk-Shiley convexo-concave mitral valves? *J Thorac Cardiovasc Surg.* 1999;117:766-74.
8. Omar RZ, Morton LS, Beirne M, Blot WJ, Lawford PV, Hose DR, et al. Outlet strut fracture of Björk-Shiley convexo-concave valves: can valve manufacturing characteristics explain the risk? *J Thorac Cardiovasc Surg.* 2001;121:1143-9.
9. Pérez A, Dennis RJ, Gil JFA, Rondón MA, López A. Use of the mean, hot deck and multiple imputation techniques to predict outcome in intensive care unit patients in Colombia. *Stat Med.* 2002;21:3885-96.
10. Altman D. Practical statistics in medical research. London: Chapman and Hall; 1991.
11. Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: application to the Cox regression model. *Stat Med.* 1992;11:2093-109.
12. Kallewaard M, Algra A, Defauw J, Grobbee D, Van der Graaf Y. Likelihood of underreporting of outlet strut fracture from examination of the Dutch Björk-Shiley CC cohort. *Am J Cardiol.* 1998;82:768-73.
13. Morton L, Omar RZ, Carroll S, Beirne M, Halliday D, Taylor KM. Incomplete and inaccurate death certification: impact on research. *J Public Health Med.* 2000;22:133-7.