April 09, 2004

MEMORANDUM TO: Nilesh C. Chokshi, Chief
Materials Engineering Branch

FROM: Mark T. Kirk, Senior Materials Engineer ***/RA/***
Materials Engineering Branch

SUBJECT: DAVIS BESSE EXPERT PANEL REVIEW

Per your direction, I have convened an independent panel to review our work assessing the structural integrity of Davis Besse. The following individuals comprised the review panel:

- Dr. William Shack of Argonne National Laboratory and the ACRS. Dr. Shack has expertise in materials analysis and corrosion.
- Dr. Gery Wilkowski of the Engineering Mechanics Corporation of Columbus. Dr. Wilkowski has expertise in fracture testing of both laboratory test specimens and large structural components.
- Professor James Joyce of the United States Naval Academy. Professor Joyce has expertise in fracture analysis and testing.

The review panel met with at ORNL in early December 2003 and had several discussions with the staff after that time. Each reviewer submitted an independent letter to the staff, but all reviewers raised the following similar themes:

- While the burst tests provide useful information on the failure characteristics of the cladding, they should not be taken to represent the conditions that existed at Davis Besse. For this reason estimates of the Davis Besse reserve capacity should be based on a finite element analysis that represents much more closely the geometric conditions that existed at Davis Besse on February 16, 2002 combined with laboratory data on strength, toughness, and failure characteristics of the stainless steel cladding.
- The burst tests require more instrumentation to permit differentiation of crack initiation and failure.
- A better characterization of crack network that existed in the Davis Besse cladding is needed to support a realistic assessment of that condition.
- Evidence on the fracture morphology of the cladding cracks does not suggest that failure was imminent on February 16, 2002.

Attached to this memorandum are the detailed letters we received from each reviewer.

Please contact me if you need any additional information or clarification.

Attachments: As stated

DISTRIBUTION: MEB DET r/f
DOCUMENT NAME: Kirk/db nilesh memo/kirk to chokshi peer review comments.wpd
OAR in ADAMS? (Y or N) Y ADAMS ACCESSION NO:__________ TEMPLATE NO. RES:______
Publicly Available? (Y or N) Y DATE OF RELEASE TO PUBLIC Immedoate SENSITIVE?________

To receive a copy of this document, indicate in the box: "**C**" = Copy without attachment/enclosure "**E**" = Copy with attachment/enclosure "**N**" = No copy

OFFICE	RES/MEB		RES/MEB					
NAME	M. Kirk */RA/*		A. Hiser */RA/*					
DATE	04/09/04		04/09/04		/ /04		/ /04	

December 12, 2003

From: Professor James A. Joyce, US Naval Academy <jaj@usna.edu>
To: Mark Kirk US NRC

Subj: Davis Besse Review Panel Initial Observations

1. I first present overall comments that came to me during the Review Panel Meeting at ORNL on Dec. 10, 2004. These observations are preliminary, and further investigations and discussions might change my opinion on many of these topics and suggestions.

2. I find it very troubling that the flaws in the Davis Besse clad material were not discovered and investigated much earlier in the investigation process. These surface flaws are apparent on a low magnification photograph, and I think they should have been observed earlier in the investigative process.

3. The disk specimen thickness should be better characterized and mapped, rather than being taken only as the result of the machining process. This should be done between the machining and testing steps on each specimen.

4. The disk specimens are just specimens and failure pressures obtained from the disk-burst experiments should not be directly related to the Davis Besse operating pressure. While it is important for these experiments to model the material, scale, temperature and loading conditions of the application, the boundary conditions are different, and the flaw geometry is different, and the failure pressures of the disk tests should not be directly related to plant operating conditions.

5. I feel that attempts to set up experiments to simulate in a real fashion the Davis Besse conditions would be extremely difficult and in all likelihood, impossible. I do think the disk specimen tests should fall within the bounds expected for the analysis, and that is not presently the case. The analysis needs to be adjusted or corrected, possible omissions explored, uncertainties need to be better quantified. I suggest the following.

6. The no-flaw disk-burst experiments It is very important to understand why the present analysis over-predicts the failure pressures of the unflawed disk burst tests. At present 1 of 2 of these tests falls below the lower 99.5% confidence bound, the second test also falls close to the lowerbound analysis prediction, and the average prediction is more than 30% below the predicted failure pressure. Major concerns here are the characterization of the PVRUF material yield stress, the disk thickness and variability, the Chakrabarty and Alexander analysis, and the burst failure criteria used. The reasons for the inaccuracy of the unflawed burst tests need to be further investigated and fully understood.

7. Flawed burst tests It was a surprise to me to see such clear evidence that the flawed burst test failures were controlled by the material fracture properties rather than by general yield and failure of the remaining uncracked ligament. This was especially surprising for the shallow crack (a/t = 0.1) disk burst test.

8. Flawed burst tests The analysis of the flawed burst tests also appears to overestimate the failure pressure by 20-30% or so. This could be due to overestimates of the burst disk thickness resulting in an overestimate of the uncracked remaining ligament, to local variations in the thickness, to overestimates of the material yield stress properties, or to overestimates of the material fracture toughness properties. Another possibility might be the effect of the biaxial loading present in the disk specimens, but not in the standard J-R specimens used to obtain the material toughness properties.

9. **Additional Tests** Tests completed to date clearly demonstrate that there are problems with the analysis, and that the problems seem to apply equally to both the unflawed disk failure predictions and the flawed specimen failure predictions. For this reason I am suspicious of the input data used for both analyses – rather then the methods of the analyses. Additional tests should only be done if they involve changes in the flaw geometry or if they include upgraded instrumentation, or both.

10. **Additional Tests** Sectioning measurements recently available on the Davis Besse clad material appear to show that the flaws present were reasonably close to the 0.05 inch depth used for the disk burst tests but did not approach the 2 inch length. Modest reductions in crack length are not likely to affect the test results, or the predictive accuracy of the analysis, but a reduction of crack length to a "penny shaped" elliptical geometry could give important information applicable to the Davis Besse analysis. I would suggest considering two tests with crack lengths of 0.5 inches and a/t = 0.2 and a/t = 0.4. The 0.5 inch length appears to be the shortest crack length observed for major flaws in the Davis Besse membrane, while the two different a/t sizes would allow a determination of whether fracture process or generally ligament yield collapse was responsible for the disk burst failure.

11. **Additional Tests** It would be very useful if any additional tests included a crack mouth opening displacement measurement (CMOD) obtained at the crack centerline during test. This is difficult because the temperature of 600^{o}F is beyond that at which standard transducers are available to make this measurement. A high temperature gage to make this measurement should again be sought. If such a gage cannot be procured running a test at a lower temperature should be considered to obtain one data set with CMOD that can be compared with computations to verify the accuracy of model measurements of this important quantity.

12. **Additional Tests** The fracture failure model includes provisions for failure at crack initiation and due to instability during ductile tearing crack extension. The loading is "load controlled" but the material strain hardens rapidly so it is not apparent whether burst failure corresponds to initiation or to a tearing instability. Much larger scatter in the failure pressure would be expected if the failure is "instability controlled" rather than "initiation controlled". A burst test that was not run to failure but stopped just before failure could provide valuable information on this issue. Two tests(#5 and #6) presently exist at a = 2 inches, a/t = 0.2. These tests show considerable scatter. Adding a third test at these conditions, to a stopping point between the failure pressures of these two tests, could provide important information on this issue, and if burst occurred, it would at least give a data set of 3 repeats for statistical scatter estimation.

13. **Additional Tests** I do not feel that there is much to be gained at present by doing a disk burst test with an array of artificial flaws. However, see 15 and 16 below.

14. **Davis Besse Issues** The clad "membrane" was measured to have bulged by approximately 0.2 inches. An analysis that models the measured shape of the clad and surrounding pressure vessel should be completed to determine if this bulge is as would be expected using the material properties of the clad and pressure vessel.

15. **Davis Besse Issues** Sections of the Davis Besse clad at FENOC have shown a complex array of flaws resulting from the corrosive etching of the ferrite component from the stainless clad material. The presence of this array of flaws could conceivably change the behavior of the clad material, increasing the resulting budge and reducing the structural capacity of the clad to contain the operating pressure. This behavior xould be very different than that resulting from a single, possible much longer flaw as used in the disk burst tests. It would be of interest to etch similar clad material, like that from the PVRUF vessel, to obtain a similar pattern of interdendritic cracks. If such damage can be simulated, it would be of interest to do a disk-burst test of this structure for comparison with tests done using artificial flaw geometries. I feel that this possibility should be investigated before a disk burst experiment is conducted on a sample containing an array of artificially flaws.

16. **Davis Besse Issues** Macro-section cladding crack observations made by FENOC show cracks that appear to have been opened by the physical loading, not just formed and extended by the preferential dissolution of the ferrite. Facing surfaces seem to show matching details implying that the cracks were formed by corrosion of the ferrite but opened by the pressure loading. The crack tips, however, appear to be sharp, not showing a crack tip opening displacement (CTOD) that is anywhere near what would be required to extend the crack by ductile tearing. Put as directly as I can, it does not appear from preliminary observations of cladding cracks on the cladding sections that the Davis Besse clad "membrane" was near to failure by fracture.

17. **Davis Besse Issues** It would be very useful to determine if the corrosive attack on the Davis Besse clad had self-arrested. Macro-section observations show a tendency for the cracks to extend so as to leave a somewhat constant remaining ligament. It appears possible that the ferrite dissolution arrested when the temperature approached some magnitude that drove out the liquid necessary for the corrosive activity to exist.

>>> Bill Shack <wjshack@anl.gov> 12/11/03 07:07PM >>>

It was a very good visit to ORNL and I would like to commend you and the ORNL staff for well prepared presentations and reviews of the analysis and tests.

The biggest puzzles in the test program for me are (1) the apparent sensitivity of what would seem to be a highly ductile material to relatively small flaws and (2) the failure of analyses to accurately predict the plastic instability of the unflawed disks.

A couple of questions follow that are applicable to both. Considering the instability problem for the unflawed specimens first:

Is the theory adequate?

The answer would seem to be yes. FEM analyses seem consistent with the C&A analysis. The original C&A paper does reference some supporting experimental data albeit on a rather different material. Has ORNL compared the C&A model to the PVRC disk tests that Riccardella analyzed? It would also be helpful to examine the displacement and strain predictions in the FEM analyses. The report gives only burst pressure. What was the criterion used to determine the burst pressure in the FEM analysis?

Are we using the right tensile properties?

The tensile data distributions (Table 1) included data for PVRUF and for other cladding. These would be appropriate distributions to use if we wanted to get the range of properties for all cladding materials, but may not be the appropriate distributions to use for comparisons with tests on the PVRUF cladding. From Table 1 it appears that the PVRUF has somewhat lower tensile properties than the other cladding materials. It may be useful to consider distributions based on just the PVRUF tensile data for comparison with the experiments. The two bounding cases of perfect correlation and no correlation between the yield and ultimate should be considered. If there is a significant difference, more realistic correlations can be considered.

Are there problems with the experiment?

I was surprised to hear that there is no post machining check on the thickness of the disks. This seems a fundamental property that should be directly measured, at least unless enough experience is gained to provide assurance that process controls alone are sufficient to determine the thickness reliably. Sensitivity calculations on the effect of thickness would be useful if they have not already been done.

For the flawed specimens:

Is the theory adequate? Different than most of the current applications of EPFM. The FM jocks will address.

Are we using the right fracture toughness properties? None of the data in the report are on PVRUF cladding. We discussed toughness data determined for PVRUF cladding, but were not shown how it fits in the distributions based on other cladding materials. Similar to the use of tensile data in the instability case, for the analysis of the experiments it may be worthwhile to consider distributions based just on PVRUF fracture toughness properties. This ought to be simple and should be considered first before trying to do experiments on constraint effects.

Are there problems with the experiment? Again it would seem verification of thickness is critical. Especially in the area where the flaw is placed.

Extrapolation of experiments and analysis to Davis Besse

Thus far everything in FM analysis applies only to experiments. Equivalent area arguments are based on instability analysis. It is important to know how sensitive the FM results are to the detailed geometry of the cavity. Calculations for some crack geometries already performed for the disk cask should be done for a different geometry with the same area. I don't know that this needs to be done for a shape that is very similar to the DB case. An ellipse of similar aspect ratio might be sufficient to determine whether the results are more sensitive to shape.

To extend the results to DB we also need to know crack geometries. Based on the overview of the failure analysis results we heard today, it appears that the 2-in. 0.05 crack is an upper bound crack size for DB. It also seems highly likely that a 0.375, 0.05 crack is a lower bound crack size. Assuming that the disk solutions are generally applicable to the DB case, the available results suggest that the upper bound crack gives unacceptable results, while the lower bound crack size will produce acceptable results. The current results also show that there is little dependence on length for cracks 1 in. or longer. Thus there is a fairly sharp transition somewhere between 0.375 and 1 in. To go further more detailed information on crack geometries is needed. Based on Mark's presentation, I doubt that this can be done with currently available information although clearly more review of the BWXT failure analysis should be done. The next step seems to be to get samples from DB for further characterization.

Next experiments

Variability in test two replicate tests performed already is fairly high, but seems consistent in variability in tensile tests for PVRUF material. Thickness variability could contribute. Analyses using just PVRUF properties data should give better indication of variability attributable to material inhomogeniety.

I would hold off on any more experiments until we can be sure that thickness is well characterized. What are new tests supposed to do?

(a) Demonstrate that strong effect of flaws is real. Benchmark analyses. Do essentially replicate tests with crack lengths of 0.375 and 1 in. Measure bulging to get something that can be compared to analysis that is less sensitive to details of exactly what constitutes failure in the model than pressure may be.

(b) Add instrumentation to get additional understanding of phenomenon. These type of tests are not my thing, but I would guess that even if we back off on temperature, it is going to be expensive and time consuming to add the additional instrumentation.

(c) Simulate DB. Rather than try to deal with interaction effects of separated cracks, try to create something that is a plausible simulation of DB. Despite the drawbacks of this approach I think it may have to be eventually done.

My priorities would be first the analyses outlined above dealing with trying to use properties that more closely match the experiment and in determining how closely you have to match cavity geometry if the mechanism of failure is ductile tearing. Then the characterization of the DB cracks. I think you will eventually have to do a (c) test. How many of the others you do depends on how deeply you feel you need to address the problem. Except for the failure to better characterize the thickness, I don't see a potential for systematic bias towards low failure pressures.