SAND2015-3360R

Report on Review of Waste Package Reliability Estimates for Geologic Disposal

Fuel Cycle Research & Development

Prepared for
U.S. Department of Energy
Used Fuel Disposition Campaign

Katrina M. Groth, Francis P. Hannigan, Huafei Liao
and Timothy A. Wheeler
Sandia National Laboratories

April 2015
FCRD-UFD-2015-000714

Revision History

Version	Description
FCRD-UFD-2015-000714	Initial issue. Assigned report number SAND2015-**** (15April2015)

Reviewed and approved for Unclassified, Unlimited Release

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy National Nuclear Security Administration under contract DE-AC04-94AL85000.

CONTENTS

FIGURES

TABLES

ACRONYMS

ASRS	Aviation Safety Reporting System
ATHEANA	A Technique for Human Error Analysis
DOE	Department of Energy
DPC	Dual Purpose Canister
EBS	Engineered Barrier System
ET	Event Tree
FEP	Feature, Event or Process
FT	Fault Tree
FV	Fussell Vesely (importance measure)
HEP	Human Error Probability
HRA	Human Reliability Analysis
IDHEAS	Integrated Decision-Tree Human Event Analysis System
IRS	International Reporting System
LPB	Low-Plasticity Burnishing
NAS	National Aviation System
NRC	U.S. Nuclear Regulatory Commission
PA	Performance Assessment
PRA	Probabilistic Risk Assessment
R&D	Research and Development
RAW	Risk Achievement Worth
RII	Risk Increase Interval
RIR	Risk Increase Ratio
RRI	Risk Reduction Interval
RRR	Risk Reduction Ratio
RRW	Risk Reduction Worth
SAPHIRE	Systems Analysis Programs for Hands-on Integrated Reliability Evaluations
SNL	Sandia National Laboratories
THERP	Technique for Human Error Rate Prediction
UFD	Used Fuel Disposition
WP	Waste Package

This page intentionally left blank.

REPORT ON REVIEW OF WASTE PACKAGE RELIABILITY ESTIMATES FOR GEOLOGIC DISPOSAL

1. Introduction

Disposal overpacks are proposed as an element of the engineered barrier system for direct disposal of spent nuclear fuel in dual-purpose canisters (DPCs) [1]. DPCs are currently licensed for storage and transport, but not disposal. In the DPC disposal system, overpacks would provide long-term containment, and conversely, they would keep groundwater from flooding DPCs. Without flooding, DPCs can never achieve nuclear criticality because they are under-moderated.

For waste isolation the overpack would be one of several barriers such as the waste form, buffer or backfill material, seals, and the natural hydrogeologic setting. For criticality control, breach of the overpack for any reason could admit groundwater, leading to degradation of aluminum-based neutron absorbing materials, and criticality. Causes of waste package (or overpack) breach include corrosion, disruptive events (e.g., seismic ground motion or human intrusion), and "early failure" due to manufacturing defects. Corrosion and disruptive events could be controlled with material and site selection, however, manufacturing defects are pervasive consequences of limited human reliability. A small enough probability of "early failure" could significantly improve the possibility of excluding postclosure criticality from performance assessment (PA), on the basis of low probability.

This report reviews previous work on "early failure" due to manufacturing defects, to project the attainable reliability of disposal overpacks for direct disposal of DPCs. It recommends approaches to inspection and analysis that could be used to improve the estimated reliability. The overall goal is to decrease the estimated unreliability of disposal overpacks below the probability screening threshold for exclusion of low probability features, events, and processes (FEPs) from further consideration.

1.1 Purpose of Report

The purpose of this report is to recommend how estimated containment reliability for waste packaging could be improved. We review an existing, previous reliability analysis for waste packages and drip shields proposed for use in a repository at Yucca Mountain, Nevada. The review identifies which failure mechanisms dominate the results and provides suggestions for improving the design, inspection, and fabrication of disposal overpacks and for updating the methodology used to estimate reliability. The work is intended to support investigations into the feasibility of disposing of spent fuel in DPCs as part of the DOE's Used Fuel Disposition (UFD) program goals.

1.2 Approach

The present study involved a two-pronged approach. Firstly, to quantitatively identify key reliability drivers in the previous work, we analyzed waste package reliability analysis developed as input to the Yucca Mountain repository license application ([10] and digital files associated with Data Tracking Numbers MO0701PASHIELD.000 and MO0705EARLYEND.000). Secondly, we leveraged Sandia experts in reliability engineering, system modeling, and human factors to identify methods to address the top drivers of unreliability.

2. Overview of Waste Packages

2.1 Physical System Representation

The general disposal concept for nuclear waste involves multiple, natural and engineered barriers. Waste packages are one element of the engineered barrier system (EBS) (Figure 1). A principal function of a waste package is to contain the waste and to isolate radionuclides from the environment. A corollary function is exclusion of groundwater that could lead to package degradation and criticality.

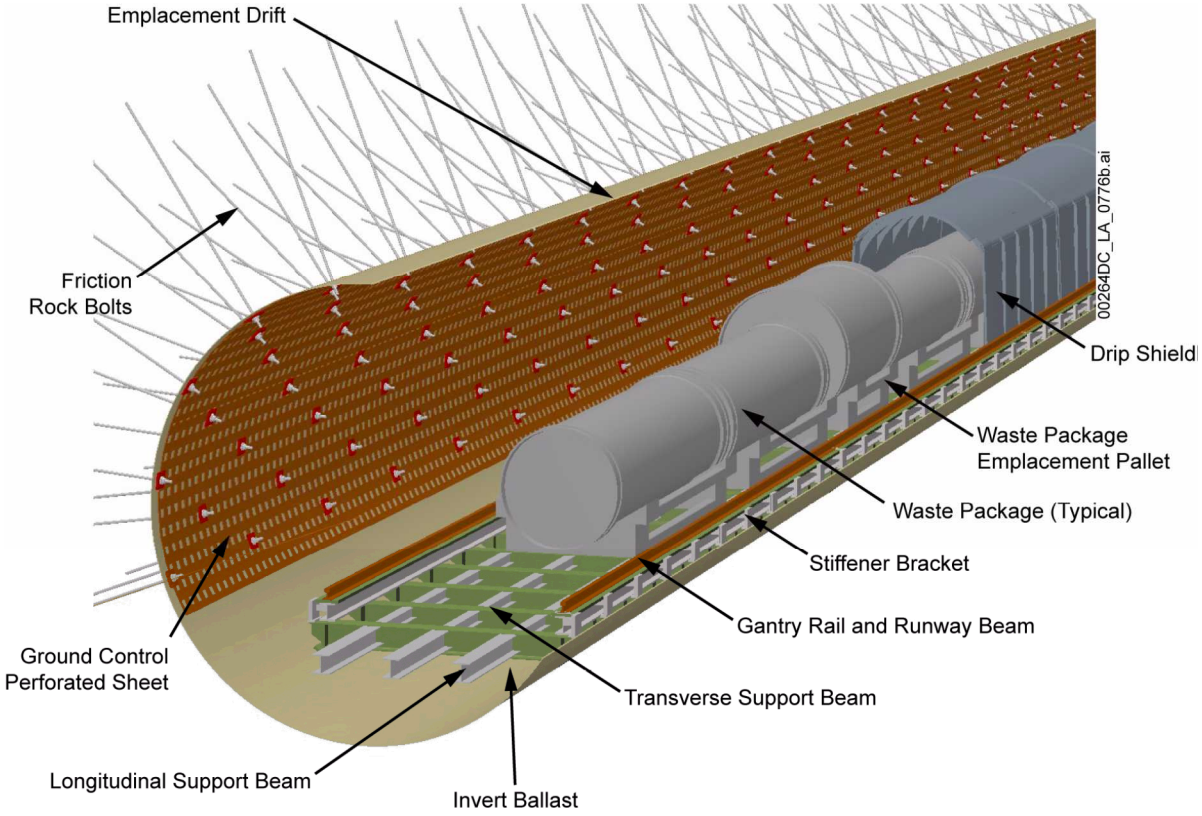

Figure 1. Schematic of an emplacement drift showing waste packages and drip shields [4]

Concepts for direct disposal of spent nuclear fuel in DPCs typically include a corrosion resistant overpack [1]. It would have two final, welded closures and would be designed for containment lifetime of 10^5 years or longer. Such waste packages would have capability to exclude groundwater well beyond the regulatory period of concern for postclosure criticality (e.g., well beyond 10,000 years).

This report addresses the reliability of these corrosion resistant overpacks, which may not perform as designed because of defects in manufacturing. Defects could result from material defects or flaws, defective welding processes, damage from rough handling, etc., and may be associated with failures in testing or inspection.

For some concepts, notably the salt repository, overpacks could be made from low-alloy carbon steel [1]. Carbon steel overpacks would have somewhat different functions (e.g., shorter containment lifetime) and are outside the scope of this report.

Many factors could affect the performance of the overpack and DPC in the event of overpack breach. These include the materials, the design of the repository and other engineered barriers, etc. While the insights in this report should hold for most corrosion-resistant overpack designs, additional analysis should be undertaken when a specific EBS design and repository site-specific information are available.

2.2 Performance Target and Failure Definition

Features, events and processes (FEPs) were developed and screened for the Yucca Mountain performance assessment [2], resulting in the inclusion of four processes/events representing modes of waste package breach: general corrosion (FEP 2.1.03.01.0A), stress corrosion cracking (FEP 2.1.03.02.0A), localized corrosion (2.1.03.03.0A), and early failure (FEP 2.1.03.08.0A). The early failure mode was incorporated probabilistically into every realization of nominal (non-disrupted) repository performance. "Failure" or breach was not described in detail but assumed to result in total loss of container functions, and to occur at zero time (i.e., at repository closure). The other three corrosion processes were modeled separately from early failure, using detailed physical process models to calculate time-dependent degradation.

By regulation [5, 17] a FEP may be excluded from post-closure performance assessment if it meets one of several screening criteria. For consideration of overpack reliability, the applicable criterion is that the predicted mean probability must be less than 10^{-8} events per year. The aggregated probability of early failure for the Yucca Mountain waste package [10] was estimated to be approximately 10^{-4} per package, which averages to approximately 10^{-4} failures per year in a repository (for 10,000 packages). This estimate is greater than 10^{-8}, so the consequences were included in the PA as noted above. Criticality was excluded because of other mitigating features such as long-lived neutron absorber plates [2]. For disposal concepts in which overpack breach is likely to cause criticality of spent nuclear fuel, exclusion of early failure could eliminate the need for mechanistic modeling of criticality events and their consequences.

If the predicted mean probability of early failure ($\bar{p}_{defect}$) is less than 10^{-8} per repository per year (integrated over all disposal overpacks), then the FEP can be excluded. Larger probabilities cannot be used for FEP exclusion, but could help to establish low risk associated with criticality events, whether they are included in performance assessment or excluded on the basis of insignificant consequences. This report addresses only the probability of early failure and not the consequences if early failure cannot be excluded on low probability.

The concept of early failure is intended to capture manufacturing and handling defects that may result in excessive degradation. It is defined as through-wall penetration of a waste package (i.e., disposal overpack) due to manufacturing or handling-induced defects, at a time earlier than would be predicted by mechanistic degradation models for defect-free waste packages. Early failure does not mean that the waste package has lost its containment function at the time of emplacement in a repository, or at the time of repository closure. Rather, breach only results from degradation processes that must still take place over time, which may extend to thousands of years or longer.

2.3 Previous Modeling Work for Waste Package Early Failure

Early failure due to manufacturing defects was analyzed using a traditional PRA approach with event trees (ETs) to model failure scenarios, and fault trees (FTs) to model root causes of failure [10]. In total, 13 defect causes were documented: weld flaws, base metal flaws, improper weld filler material, improper stress relief for lid (low plasticity burnishing), improper heat treatment, improper weld-flux material, poor weld-joint design, contaminants, improperly located welds, missing welds, handling-induced defects, emplacement errors, and administrative or operational errors. Six of those causes were screened out on low probability or low consequences. Administrative or operational errors were treated as contributors to the other failure mechanisms.

The remaining seven (heat treatment is performed independently for two components) that were analyzed for waste package early failure are:

1) Weld flaws (ref. [10], Section 6.3.1)
2) Improper base metal selection (ref. [10], Section 6.3.2)
3) Improper heat treatment of outer corrosion barrier shell (ref. [10], Section 6.3.3)
4) Improper heat treatment of outer corrosion barrier lid (ref. [10], Section 6.3.4)
5) Improper stress relief of outer corrosion barrier lid (low plasticity burnishing) (ref. [10], Section 6.3.5)
6) Waste package mishandling damage (ref. [10], Section 6.3.6)
7) Improper weld filler material. (ref. [10], Section 6.3.7)

Six of these causes (items 2 through 7) were modeled via event trees and fault trees. Weld flaws were modeled separately via physical models; the results are not combined with the other early failure causes. Each of the six failure mechanisms was modeled as the initiating event in an ET. For all ETs, the target end state is "DAMAGED-WP." The ETs also contain one to three events that would identify the occurrence of the failure mechanism and lead to a "REJECTED-WP" end state. Figure 2 shows one of the ETs from the previous analysis and Table 1 summarizes the level of detail seen in all six models. As can be seen, the failure mechanism is decomposed at a very high level. Both pivotal events in this tree are directly assigned probabilities of occurrence (i.e., they do not have associated FTs). The most complex ETs from this analysis contain four pivotal events after occurrence of the failure mechanism. The pivotal events and basic events in the ETs and FTs are largely human-caused. See Appendix A for a complete list of basic events and basic event probabilities included in the analysis. Most of these events were quantified with the Technique for Human Error Rate Prediction (THERP) methodology.

The calculations performed in the previous analysis were replicated in this review, and were found to be performed correctly. However, the level of decomposition of the manufacturing process is quite general and typical of screening-level analyses. The failure mechanisms considered are mostly human failure events with one human-driven opportunity for recovery. There are no events in the model that credit engineered systems that are designed to prevent or mitigate the effects of human errors.

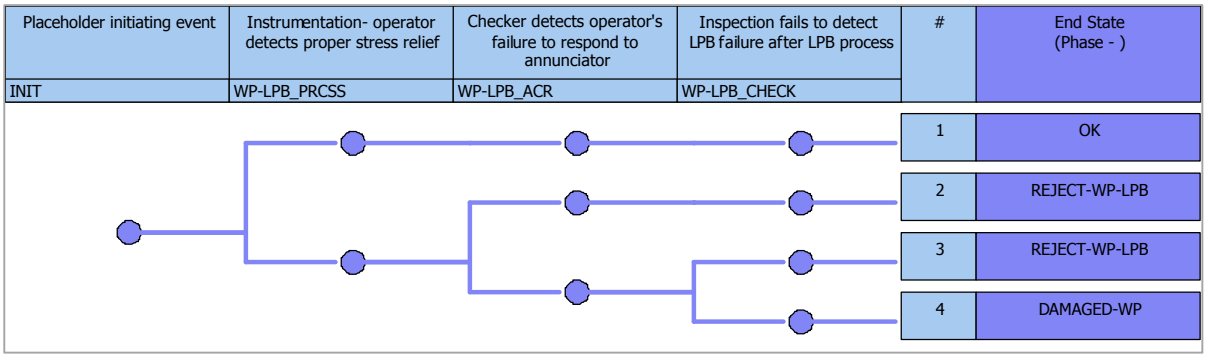

Figure 2. Representative event tree from 0705EARLYEND analysis showing level of decomposition for the occurrence of the low plasticity burnishing failure mechanism.

Table 1. List of event trees in the 0705EARLYEND 2009 analysis.

Number of >>>	Top Events in ET	End States in ET	Damaged WP End States	FTs	Directly Input Probabilities	BEs in FTs	Distributions Used for Probabilities
Improper base metal selection	2	3	1	0	2	--	Lognormal
Improper heat treatment of outer corrosion barrier	5	6	2	4	1	2, 2, 2, 3	Lognormal
Improper heat treatment of outer corrosion barrier lid	5	6	2	3	2	2, 2, 3	
Improper stress relief of outer corrosion barrier lid (low plasticity burnishing)	3	4	1	1	2	3	Lognormal
Waste package mishandling damage	2	3	1	1	1	8	Lognormal
Improper weld filler material	2	4	1	0	2	--	Lognormal
ET = event tree; WP = waste package; FT = fault tree; BE = basic event							

Table 2. Event tree end state results for the six failure mechanisms analyzed in SAPHIRE.

Event Tree	Probability in SAPHIRE (mean)	Probability in SAPHIRE (point estimate)	Probability in SAPHIRE (median)
Base metal flaw	1.251e-7	1.25e-7	7.960e-8
Heat treatment shell	3.726e-5	3.234e-5	1.423e-5
Heat treatment lid	3.497e-5	3.106e-5	1.354e-5
Low plasticity burnish	3.769e-5	4.290e-5	7.274e-6
Mishandling	9.708e-7	9.60e-7	2.857e-7
Weld filler flaws	1.251e-7	1.25e-7	7.960e-8
Total Damaged-WP (Sum of above)	**1.11E-04**	**1.075e-04**	**3.55E-05**

SAPHIRE 8 [25] is a probabilistic risk and reliability assessment software tool. SAPHIRE stands for Systems Analysis Programs for Hands-on Integrated Reliability Evaluations. The system was developed for the U.S. Nuclear Regulatory Commission (NRC) by the Idaho National Laboratory. Table 2 contains the results of the solution and probabilistic quantification of the six event trees documented in the 0705EARLEND 2009 analysis. Three statistics are shown for each fault tree top event: the mean, the point estimate calculation (no probability distributions were sampled to account for the aleatory uncertainty of the fault tree basic events), and the median.

Note that three processes (heat treatment of the shell, heat treatment of the lid, and low plasticity burnishing) have failure probabilities roughly two orders of magnitude greater than the remaining four.

2.4 Models for Waste Package Corrosion Failures

The drip shields and waste packages comprise a defense-in-depth approach to protection of the waste, particularly with respect to localized corrosion of the waste package. Localized corrosion of the waste package is only possible if the drip shield fails to perform its function. Such a failure of the drip shield would allow incoming seepage to contact the waste package and if certain aggressive environments are present. The treatment of the corrosion processes that may affect the waste package is discussed below.

General Corrosion of Waste Packages—General corrosion rates of the waste package's Alloy 22 are sufficiently low that the waste packages will last for long periods of time. This is an important feature of the waste package performance as long as the drip shield performs its function to protect the waste package from seepage.

Localized Corrosion of Waste Packages—Localized corrosion mechanisms on the waste package surface are dependent upon the thermal and chemical environment on the waste package surface. Localized corrosion would only occur if the drip shield fails to perform its function, whether the drip shield fails due to corrosion or to early failure. Localized corrosion of the waste package is modeled as a degradation mechanism for those cases where the drip shield fails.

Stress Corrosion Cracking of Waste Packages— As stated in section 2.2 stress corrosion cracking is the initiation and propagation of cracks due to three simultaneous conditions: a susceptible material, critical environment, and sustained tensile stress. Stress corrosion cracking of Alloy 22 is modeled to occur as a result of mechanical degradation following seismic events and in the closure weld lid region for the nominal scenario. Such stress cracks are sufficiently small and tight to allow only diffusive transport of radionuclides through the cracks. Stress corrosion cracking models and data are presented in [10], Section 2.3.6.5.

2.5 Identification of Top Unreliability Drivers for Early Failure Events

2.5.1 Importance Analysis Methodology

Importance measures provide quantitative means for understanding how model parameters (e.g., failure mechanisms) affect reliability. These importance measures can provide insight into which components dominate the reliability calculation, which components are safety critical, and how much reliability improvement could result from significant changes in a single component. In

general, the best practice is to use one risk-reduction focused measure, and one risk-increased focused measure, and compare results of both analyses.

Four traditional importance measures are available in SAPHIRE: Birnbaum, Fussell Vesely (FV), Risk Reduction Ratio or Interval (RRR/RRI), and Risk Increase Ratio or Interval (RIR/RII). For this review we choose to discuss the RRR/RRI and RIR/RII, which are defined below. The FV and Birnbaum are defined in Appendix A. For all the importance measures:

- $F(x)$ is the original probability of the end state ($p\{\text{Damaged-WP}\}$)
- $F(0)$ is the probability of the end state with the event probability set to 0 (perfectly reliable).
- $F(1)$ is the probability of the end state with the event probability set to 1 (failed)

Risk Reduction Ratio (RRR) or Interval (RRI) are two related measures expressing how much risk would decrease if the basic event probability is zero (i.e., the component is perfect and never fails). Both measures are also referred to as risk reduction worth (RRW). RRI is also called inspection importance, because it denotes which components are most important to inspect.

$$RRR = F(x) / F(0)$$

$$RRI = F(x) - F(0)$$

When RRR equals 1.0, there is no reduction in risk. Larger RRR values indicate larger decreases in risk if the component is made more reliable.

Risk Increase Ratio (RIR) or Interval (RII) are two related measures expressing how much risk would increase if the basic event probability is equal to 1 (i.e., the component fails). Both measures are also referred to as RAW (Risk Achievement Worth).

$$RIR = F(1) / F(x)$$

$$RII = F(1) - F(x)$$

When RIR equals 1.0, the risk stays the same. Larger RIR values denote larger increases in risk if the component fails.

2.5.2 Results: Top Drivers of Canister Early Failure

Using SAPHIRE, we conducted an importance analysis on the DAMAGED-WP end state from the previous early failure analysis [10]. This section presents the results for the five most important basic events for both types of importance measures. In total, eight basic events are discussed in this section. Full importance measure results are presented in Appendix A.

Two processes dominate the unreliability of the waste package: heat treatment and low plasticity burnishing (LPB); this is seen in both the event tree results in Table 2 and again in the importance measure results in Table 3.

Table 3. Importance measure results for the top drivers of canister early failure (values for the top drivers are in bold).

Name	#	Prob.	RRR (RRI)	RIR (RII)	Description

Name	#	Prob.	RRR (RRI)	RIR (RII)	Description
HT_OPERATOR_ERROR	7	3.00E-03	**1.81 (4.81E-05)**	**149 (1.59E-02**)	Heat treatment process operator fails to respond to alarm
LPB_CHECK	3	1.60E-01	**1.66 (4.29E-05)**	3.10 (2.25E-04)	Inspection fails to detect LPB failure after LPB process
LPB_ACR	3	8.10E-02	**1.66 (4.29E-05)**	5.53 (4.87E-04)	Checker detects operator's failure to respond to annunciator
LPB-OPERATOR	1	3.00E-03	**1.57 (3.89E-05)**	**121 (1.29E-02)**	Operator fails to responds to annunciator
HT_INSPECT	6	1.60E-01	**1.18 (1.66E-05)**	1.81 (8.71E-05)	Inspection detects improper heat treatment cooldown?
LPB-IC	1	3.00E-04	1.04 (3.89E-06)	**122 (1.30E-02)**	Instrumentation & control system fails to alarm
TIMER_FAILURE	4	1.20E-04	1.02 (1.80E-06)	**140 (1.49E-02)**	Timer alarm fails to alarm
LPB-SENSOR	1	1.00E-05	1.00 (1.30E-07)	**122 (1.30E-02)**	Pressure monitor to LPB hydraulic system fails

Based on the RRR, the five events that would most improve the reliability of the canister are: HT_OPERATOR_ERROR, LPB_CHECK, LPB_ACT, LPB-OPERATOR, and HT-INSPECT. These five events are the top candidates for changes, because they will have the greatest impact on the mean probability of early failure.

Unfortunately, the RRR results indicate that no single change to the system would drive disposal overpack unreliability below 10^{-8} failures/year. The highest RRR in the early failure analysis is 1.81, for HT_OPERATOR_ERROR (heat treatment process operator fails to respond to an alarm). This RRR indicates that risk could be reduced by a factor of 1.81 if the probability of HT_OPERATOR_ERROR was 0 (i.e., the operator performed perfectly). This type of change is not sufficient to reach the low-probability screening threshold for overpack early failure. Additionally, it is unlikely that any human event could be made perfect; in many human reliability analysis (HRA) methods, the lowest human error probability (HEP) available is approximately 10^{-5} to 10^{-6}.

Based on the RIR results, the five events that would most reduce reliability of the canister are: HT_OPERATOR_ERROR, TIMER_FAILURE, LPB-IC and LPB-SENSOR and LPB-OPERATOR. The RIR results indicate that these five events are critical aspects of reliability assurance in the DPC process. If any of these events were removed, the mean probability of early failure would increase by two orders of magnitude.

3. Challenges, Opportunities and Research Directions

This section is broken into three parts. In Section 3.1, we discuss the challenges and conservatisms associated with the previous work. In Section 3.2 we discuss opportunities for improvement of the analysis which may provide near-term improvements in the analysis of disposal overpack reliability by reducing certain conservatisms discussed in Section 3.1.1. In Section 3.3, we discuss research directions that could improve canister reliability and its estimation, and address conservatisms in previous work. These research directions may also have application beyond DPC disposal overpacks.

3.1 Challenges

3.1.1 Key Conservatisms

The development of both the performance of the EBS in the repository environment with regard to corrosive processes, and the development of the early failure model contain numerous modeling conservatisms. For example:

- **Immediate, complete failure of affected waste packages and drip shields** – Any waste package (or drip shield) that exhibits early failure is modeled as completely failed at repository closure (time = 0). In other words, the waste package (or drip shield) functions are completely nullified for early failures. This ignores environmental effects such as the dependence of corrosion on evolving humidity conditions in the repository environment. It also ignores the particular characteristics of corrosion failures, such as the potentially limited extent of localized corrosion or stress corrosion cracking (which would not affect the entire package surface). Because the long-term corrosion performance of "early-failed" components was not modeled, the extent to which this major conservatism hastens or overestimates the occurrence of conditions favorable to a criticality is not known. Regardless, it is intuitive that early-failed waste packages (and drip shields) would continue to perform their containment functions for some time before penetration were to actually occur, possibly beyond the 10,000-year period of concern for FEP screening.
- **Stress corrosion initiation from weld flaws -** Weld flaws in the closure-lid weld are possible sites for the initiation of stress corrosion crack growth. Data for calculating the probability of non-detection of weld flaws are based on experiments performed in the late 1970s. Detection capabilities have likely improved since then.

There are numerous modeling conservatisms in this early failure analysis, and several are discussed in greater detail in Sections 3.1.2 through 3.1.4 of this report. The PRA model used to calculate the early-failure probabilities uses a simplistic representation of the capability to detect defects and to initiate recovery or restorative actions during the manufacturing process. There are no events in the model that credit engineered measures designed to prevent or mitigate the effects of human errors. Furthermore, the HRA method of Technique for Human Error Rate Prediction (THERP) [17], which was used in the study, was developed several decades ago and does not reflect the current state-of-the-art.

3.1.2 Limited Information About Overpack Design and Manufacturing

Available design information on the process included only high–level descriptions of the main human tasks. To understand how humans can enhance or decrease overpack reliability, it is important to understand how humans will interact with the system (this is reflected in the recommended research directions in Section 3.3).

Management of the risk associated with potential operator failure or unsafe operator actions in waste package manufacturing and handling needs to be based on what is currently known about the behavior of a plant and its operators. However, our current state of knowledge is that the process, plant design, and operation of waste package manufacturing and handling are loosely defined and characterized. More detailed information is needed to perform a detailed process and task analysis to understand the role of human operators in the process, what tasks operators need to perform, how operators may fail at the tasks, and what factors can potentially impact operator performance. Thus, the gap in our knowledge limits the ability to identify and treat human failures, including those not yet seen in actual incidents or accidents. It also limits the ability to identify effective measures to enhance human performance to reduce failure probability.

3.1.3 HRA Method Limitations and Applicability Questions

Current HRA methodologies are not ideally suited for understanding errors that can occur in design or implementation of manufacturing processes [7]. HRA for a plant in the design phase is challenging because appropriate boundary conditions for an accident event can be difficult to define. As a result, the targets and functional requirements for human factors engineering can be difficult to define. Many popular HRA methods are specifically intended to identify and quantify errors made by personnel in existing nuclear power plants. The use of HRA in the design and implementation of processes and facilities for high-reliability fabrication has received relatively little attention.

Some older HRA methods, e.g., THERP, were specifically developed for manufacturing activities but they tend toward overly simplistic description of the human cognitive processes that affect HEP estimates. This older approach to HRA is inconsistent with more recent understanding of human behavior and the relationship between that behavior and performance of an industrial system. Additionally, older HRA methods cannot accommodate recent advances such as the digital human-machine interface, automation, and software systems [7] that are very likely to be used for expensive high-reliability fabrication.

Some researchers have recently published findings critical of the use of HRA [6], including that "the state of the art in quantitative HRA is too poor to make the summative assessments of risk and reliability that our regulators assume" and that "very little use of this extensive, often empirically based literature has been made in developing HRA methodologies." However, despite the shortcomings of even modern HRA methods, they still serve an important need to consider human contributions to system failures. HRA methods remain the best way to include humans in a PRA context.

3.2 Opportunities for Improvement Within Current Models

The conservatisms discussed above are largely related to choices about the structure of the ET and FT models and also to choices about the probabilities in the models. These types of choices are not necessarily independent, and both could be addressed to model the disposal overpack

early failure probability. In this section we present three options: the first would update the HRA modeling approach, the second would refine probability estimates, and the third would enhance model structure by including additional inspection and monitoring.

3.2.1 Develop a New Early Failure PRA/HRA Model

As discussed above the previous early failure models are posed at a relatively high level, and do not include events that credit engineered systems that could prevent or mitigate the effects of human errors. Furthermore, the HRA method used (THERP) was developed several decades ago and does not reflect the current understanding of failure mechanisms and quantification of human failure in the scientific community.

A promising opportunity for improving the previous modeling approach is to develop a new PRA model at a more rigorous level of detail, using an updated HRA approach. This would provide more detailed insights into the causes of early failure and the opportunities for reducing the defect rate via both human and hardware/software interventions. Two NRC HRA methods, A Technique for Human Event Analysis (ATHEANA) [11], and Integrated Decision-Tree Human Event Analysis System (IDHEAS) [13] show such promise.

The HRA portions of the model could be improved using the ATHEANA method [12], which was developed to provide support for identifying and modeling key human/system interactions and their effects at nuclear power plants and similar facilities. ATHEANA is one of only a few modern HRA methods currently available to address the challenges discussed in Section 3.1.3; other methods are developmental or are not widely used in the U.S.

The key tasks in developing a new model for the analysis of the disposal overpack manufacturing would include:

- Inclusion of hardware and software failure and recovery opportunities in the ET and FT models.
- Quantification of hardware and software elements using available system reliability data.
- Use of ATHEANA to identify human failure events, unsafe acts, and determine why they could occur.
- Use of ATHEANA to quantify the probabilities of unsafe acts and error-forcing contexts for the specific processes relevant to early failure.
- Elicitation of quantitative information specifically about disposal overpack manufacturing and handling processes.

Ongoing work at the NRC has developed the cognitive basis for development of a science-based HRA method. This cognitive basis is being implicitly considered in development of the IDHEAS method and is targeted at nuclear power plant operations. By rooting improvements to HRA in psychological concepts rather than design specifics, the HRA could remain sufficiently technology-neutral to represent both current and future processes. This could be a significant advancement for waste disposal applications that describe manufacturing processes taking place over several decades, which are likely to use evolving technologies.

The updated model would provide more rigor in both qualitative and quantitative information. It would represents the current state of knowledge regarding: 1) what human actions are important

to risk; 2) what design, operation, and environmental aspects of the plant could significantly impact operator performance; and 3) what elements are not significant, and why. Qualitative information conveys insights into failure mechanisms, risk sources, and possible process improvements. An updated model would provide a stronger foundation for future early failure analysis, although it is unclear if the results could be used to exclude the early failure FEP.

3.2.2 Update Human Error Rates Using Newer HRA Models

The Yucca Mountain early failure analysis report [10] largely uses HRA data from THERP for human error probability quantification. More recent HRA methods such as SPAR-H [13], CREAM [14], and IDHEAS [15] would provide different human error probabilities for the basic events in the early failure models. One opportunity for improvement would be to update the probabilities using a newer HRA method. As shown in the importance analysis results in Section 2.4.2, updating any single probability in the model would not sufficiently reduce the probability of early failure. However, the use of a new HRA method would systematically change all of the probabilities. It is unclear if this change would substantially reduce the total probability of early failure.

To provide further insight, we ran an example calculation using information from the IDHEAS HRA method. In the Yucca Mountain analysis, two of the top drivers were operator errors in responding to alarms during heat treatment and low plasticity burnishing (HT_OPERATOR_ERROR and LPB-OPERATOR). Both events were assigned the probability of 3.0×10^{-3}, which may be conservative. In IDHEAS, the nominal HEP for this failure mode is 2.4×10^{-5}. We updated these probabilities and reran the calculation of the point estimate for DAMAGED-WP. The original point estimate for DAMAGED-WP was 1.075×10^{-4}. For the revised case the new point estimate for DAMAGED-WP is 2.13×10^{-5}. Additional reduction may be seen by updating the probabilities of the remaining basic events. However, some increase may also be seen if THERP values are found to be non-conservative.

It is likely that updating the human error probabilities will lead to order-of-magnitude reduction in the probability of early failure, but it is unclear if the results could be used to exclude the early failure FEP.

3.2.3 Credit Additional Inspections and/or Monitoring Systems

Another option for increasing reliability of the overpack involves crediting engineered systems that are designed to prevent or mitigate the effects of human errors, and adding additional checks or monitoring systems during critical aspects of the process (e.g., during heat treatment and low plasticity burnishing). From a reliability perspective, adding well-designed monitoring systems to the disposal overpack manufacturing process could permit significant (several orders of magnitude) reduction in the unreliability of the canisters. Realizing such gains would require substantial restructuring of the fault trees and event tress in the previous analysis.

Additional hardware or software-based monitoring systems could substantially improve the reliability of the manufacturing process, and would be preferable to human-based monitoring or checking systems.

Additional human-based systems could also improve reliability if they are appropriately designed to enable the reliability improvement. Adding human-based steps comes with all of the performance variability associated with human reliability. Research suggests that human error

probability may actually *increase* if an inspection system is implemented without regard for psychological, social, and task characteristics involved when one human inspects the work of another. Considerations include the number of checkers (two may be optimal) [18,19,20], checker independence [21], and frequent rotation of checkers to avoid routine [11,22]. Taking credit for (additional) monitoring and checking process features could reduce early failure probability by several orders of magnitude if the features are adequately designed.

3.3 Suggested Research Directions

Section 3.2 suggested available opportunities for improving models of early failure. However, these activities may not address all of the challenges discussed in Sections 3.1.2 and 3.1.3. This section describes activities that could address limitations on information (Section 3.3.1), and improve the quantification of human error probabilities for waste-specific applications (Sections 3.3.2 and 3.3.3).

3.3.1 Better Define the Effects From Human Error on Disposal Overpack Design and Manufacturing via Task Analysis and Expert Elicitation

Task analysis is concerned with determining the human, mechanical, and cognitive processes required to accomplish a particular task. This takes into account not only the actions required to complete the task, but also performance-shaping factors such as organizational and environmental influences that may have an impact on task execution.

Human performance can be modeled as five macro-cognitive functions: *detecting and noticing*, *understanding and sensemaking*, *decision making*, *action*, and *team coordination* [16]. Most operator actions modeled in the previous analysis [10] fall in the macro-cognitive functions of detecting and noticing, and action. In particular, the top five performance drivers identified above (two relate to operators failing to respond to alarms or annunciators, two relate to inspection failure, and one relates to failure of immediate peer check) can be classified as failures of detecting and noticing. Other macro-cognitive functions may also need to be modeled, for example, whether error outcomes change if operators assess plant status, make response plans, and communicate and coordinate with each other. Failures related to all macro-cognitive functions need to be identified, evaluated, and treated in the system model (e.g., event trees and fault trees) and neglect of any macro-cognitive function needs to be justified. Note that this would require detailed description of the manufacturing and handling processes.

When evaluating a failure mode, one needs to consider various contextual factors (e.g., performance shaping factors) that can impact the likelihood of the failure mode. For example, the probability of an operator failing to respond to an alarm can be influenced by many factors, including the operator's training and experience, quality of human-machine interfaces, and use of procedures.

Expert elicitation is a promising way to compensate for information insufficiency that may hinder detailed task analysis. Experience and judgment of experts from similar domains can help elucidate tasks that were not included in the previous model, identify other aspects of human performance, characterize contextual factors, and determine the applicability of knowledge from other domains. Additionally, probability distributions can be elicited from the experts to inform and calibrate HEP quantification (see Section 3.3.3).

3.3.2 Develop or Adapt HRA Methods to Disposal Overpack Manufacturing and Handling

Lack of an HRA method developed specifically for waste package manufacturing and handling calls for a new HRA method or adaptation of an existing method. One research path would be to develop a new HRA method to address the different types of activities relevant to design and manufacture of disposal overpacks. A related path would be to adapt an existing HRA method specifically to waste packaging applications. We note that using an HRA method outside the scope for which it was originally intended can lead to significant uncertainties and arbitrariness.

In both of these research approaches, the HRA method should be context-centered (i.e., specific to processes and facilities to be used in waste packaging) because human performance depends heavily on contextual factors. Furthermore, the method should have a theoretical basis to address operator cognitive failures, and account for possible development of inappropriate plant status assessments, goals, and action plans, and the tendency for operators to persist in process implementation even when status is contra-indicated by observation.

One challenge to adapting an existing HRA method is the content validity ("face" validity). That is, whether the failure modes and/or important performance shaping factors in waste package manufacturing and handling are within the scope of the method, and whether the method has appropriate structure to represent the most likely types of human failure. Content validity can be evaluated through subject matter expert or examination of related operating experience. As discussed in Section 3.2.1, ATHEANA [12] and IDHEAS [15] are methods that could be adapted to address many of the human-machine interface complexities that characterize advanced, digitally controlled processes in non-nuclear plant context.

3.3.3 Data Collection Framework for Human Performance in Disposal Overpack Manufacturing

It is widely acknowledged within the HRA community that human performance data are needed to support HRA applications and method development (see Appendix C for further discussion). An alternative to developing a comprehensive HRA methodology is to develop a framework for collecting data on human performance in the context of disposal overpack manufacturing and handling, using a voluntary safety reporting system. Data collected during operations can then be used to analyze and improve processes as needed.

Safety reporting systems are in use to varying extents and levels of success across many domains. An example is the Aviation Safety Reporting System (ASRS), a national database of safety information voluntarily submitted by pilots, cabin crew, mechanics, air traffic controllers, and others in a position to report on safety issues within the National Aviation System (NAS). The NAS has been collecting data via ASRS for the U.S. Federal Aviation Administration for 38 years. During this time the industry has built a safety culture in which employees making reports are assured that the system is confidential, voluntary, and non-punitive. Since 1988, ASRS has seen a 285% increase in submitted reports, and the system has received more than 1.1 million reports since its inception [3].

Submissions are reviewed by experts in the domain and classified for actions ranging from immediate response when safety is at risk (Alert Bulletins), to those reserved for future action in collaboration with safety officials. Sixty-four research studies and special papers have been published using ASRS data, in areas such as operations, human factors, and the utility of

confidential reporting [3]. The ASRS model is now used internationally, has been sanctioned by the International Civil Aviation Organization as the standard for voluntary reporting systems, and is currently being applied in the railroad and firefighting domains [3].

Similarly, the nuclear energy domain has built an International Reporting System (IRS) for operating experience, in which 31 member nations participate. According to the IRS website, "…participating countries exchange experience to improve the safety of nuclear power plants by submitting event reports on unusual events considered important for safety" [8]. The details of each report are not public, which is thought to promote honest and detailed submissions. Output includes generalized reports (Nuclear Power Plant Operating Experience) every 2 to 3 years, and topical studies that are conducted to "identify the main recurring cause, contributing factors, lessons learnt and to disseminate and promote recommendations aiming at reducing the reoccurrence of similar events in the future" [9].

There are important differences between the two reporting models. Where ASRS encourages voluntary reports from the workforce without regard for how large or small a condition is, or to whether an incident actually occurred, the IRS only contains reports submitted to a National IRS Coordinator who then deemed the event worthy of international consideration. The latter type of reporting precludes insight into latent or active errors that may have been corrected at some level of the process, but potentially occur every day across the industry, and will likely continue to occur without intervention, or until a major disaster occurs. The report for each reporting period, produced by the International Atomic Energy Agency, is at a high level with little actionable information.

Both of these reporting systems are used primarily in an operating environment and not for manufacturing. It is increasingly important to capture errors in manufacturing or fabrication, particularly for items that perform in high-risk, complex situations. Inevitably, some failures that occur in operating environments will have root causes in the manufacture and fabrication of system components.

Modifying the ASRS concept for use within the nuclear energy domain, particularly for manufacturing and fabrication, would improve safety. The challenge in developing such a reporting system is in creating a safety culture throughout the organization that supports this type of reporting. Condition and incident reporting systems are already parts of the nuclear safety culture fostered in operating nuclear plants [24]. Adopting nuclear safety culture in manufacturing of disposal overpacks, with a focus on continuous improvement, could accomplish the objectives of a safety reporting system.

4. Summary and Conclusions

Overpack reliability is a potentially important factor in preventing or limiting the incidence of nuclear criticality of DPC-based waste packages in certain disposal environments. Specifically, in environments for which flooding of breached waste packages with fresh water (less chloride than seawater) is possible, and in which the aluminum-based neutron absorber materials used in DPCs readily degrade.

In analysis performed to support the Yucca Mountain license application [10], waste package reliability was expressed in terms of the probability of stylized early failure, aggregated from several types of defects during manufacture. That analysis is reviewed in this report, and while the calculations were found to be performed correctly, the previous approach did not describe waste package manufacturing in much detail. The conceptual model used for human error was simplistic, with little or no recognition of the capability to detect defects and initiate recovery or restorative actions during manufacturing. No credit was allowed for engineered measures designed to prevent or mitigate the effects of human errors. Also, the THERP method [17] used in the analysis was developed several decades ago and does not reflect the current state of the art in HRA.

This report identifies and discusses several methodology options to improve the estimated reliability of disposal overpacks such as those that could be used for direct disposal of DPCs. The goal of improvement would be to decrease the probability of overpack early failure, below the probability threshold that allows exclusion of FEPs from performance assessment, and to thereby exclude postclosure criticality. The previous early failure analysis produced an aggregated mean probability of approximately 10^{-4} early failures per waste package (with approximately 10,000 waste packages). The applicable FEP screening threshold is 10^{-8} early failures per year over all waste packages, integrated over 10,000 years. For 10,000 waste packages this can be interpreted to mean a 10,000-fold improvement in overpack reliability is needed for FEP exclusion.

As discussed in Section 3.2, an enhanced manufacturing process that ensures reliability through prevention, detection, and mitigation of flaws or errors could reduce the probability of early failure by several orders of magnitude. This finding also extends to the overpack design, for example, the overpack design could include multiple, independent barriers that reduce the joint probability of failures. The potential value of enhanced design and manufacturing processes could be simulated using PRA.

The most rigorous near-term option for improving estimated reliability is development of a new PRA model for early failure, based on detailed description of the overpack design and an enhanced manufacturing process. This approach would address both the quantitative and structural issues identified with the previous waste package early failure analysis, and would be the most straightforward way to address the limitations of HRA methods that are not specifically developed for manufacturing applications.

We also propose longer-term research to examine: 1) the role of humans in overpack design and manufacturing, via task analysis and expert elicitation; and 2) developing or adapting an updated HRA method for waste management applications.

5. References

[1] Hardin, E., C. Bryan, A. Ilgen, E. Kalinina, K. Banerjee, J. Clarity, R. Howard, R. Jubin, J. Scaglione, F. Perry, L. Zheng, J. Rutqvist, J. Birkholzer, H. Greenberg, J. Carter, T. Severynse. 2014. *Investigations of Dual-Purpose Canister Direct Disposal Feasibility*. FCRD-UFD-000069 Rev. 0, Albuquerque, NM: Sandia National Laboratories

[2] U.S. Department of Energy, ANL-WIS-MD-000027 REV00, Office of Civilian Radioactive Waste Management, Las Vegas, NV, March 2008,

[3] Linda J. Connell. Aviation Safety Reporting System (ASRS) program briefing. June, 2014.

[4] DOE (U.S. Department of Energy) 2008. *Yucca Mountain Repository License Application, Safety Analysis Report*. DOE/RW-0573 Rev. 0. Office of Civilian Radioactive Waste Management. Washington, DC.

[5] EPA (U.S. Environmental Protection Agency) 2008. 40 CFR Part 197: *Public health and environmental radiation protection standards for Yucca Mountain, Nevada; final rule*. Federal Register 73(200):61256–89.

[6] French, S., T. Bedford, S.J.T. Pollard and E. Soane 2011. "Human reliability analysis: A critique and review for managers." *Safety Science*, 49(6):753–763, July 2011.

[7] Groth, K.M., J.A. Forester, H. Liao and T.A. Wheeler 2011. Applicability of human reliability analysis and identification of research needs for NGNP HRA. Letter Report JCN N6921, Task 3. Prepared for Office of Nuclear Regulatory Research, U. S. Nuclear Regulatory Commission. February, 2011.

[8] IAEA (International Atomic Energy Agency) International Reporting System for Operating Experience (IRS). Accessed January 23, 2015. http://nucleus.iaea.org/CIR/CIR/irs1.html.

[9] OECD/NEA (Organization for Economic Cooperation and Development/Nuclear Energy Agency) 2011. *Nuclear power plant operating experience from the IAEA/NEA international reporting system for operating experience 2009-2011*. Technical Report NEA 7120. December, 2012. ISBN 978-92-64-99193-4.

[10] SNL (Sandia National Laboratories) 2007. *Analysis of Mechanisms for Early Waste Package/ Drip Shield Failure*. ANL-EBS-MD-000076 REV 00. Prepared for the U.S. Department of Energy, Office of Civilian Radioactive Waste Management, Las Vegas, NV. June, 2007.

[11] Swain, A.D. and H.E. Guttman 1983. *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. NUREG/CR-1278-F, U.S. Nuclear Regulatory Commission, Washington, D.C.

[12] NRC (U.S. Nuclear Regulatory Commission) 2000. *Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA)*. NUREG-1624, Rev. 1. Washington, D.C. May, 2000.

[13] Gertman, D., H. Blackman, J. Marble, J. Byers, L. Haney and C. Smith 2005. *The SPAR-H Human Reliability Analysis Method*. NUREG/CR-6883. U.S. Nuclear Regulatory Commission, Washington, D.C.

[14] Hollnagel, E. 1998. *Cognitive reliability and error analysis method CREAM*. Elsevier Science Ltd.

[15] NRC (U.S. Nuclear Regulatory Commission) 2012. *NRC/EPRI Draft Report on An Integrated Decision-Tree Human Event Analysis System (IDHEAS) Method for NPP Internal At-Power Operation*. Washington, D.C.

[16] Whaley A.M., J. Xing, R.L. Boring, S.M.L. Hendrickson, J.C. Joe, K.L. LeBlanc and E. Lois 2012. Building a Psychological Foundation for Human Reliability Analysis. NUREG-2114. U.S. Nuclear Regulatory Commission, Washington, D.C.

[17] NRC (U.S. Nuclear Regulatory Commission) 2003. *Yucca Mountain Review Plan: Final Report*. NUREG-1804. Washington, D.C.

[18] Wiener, E. L. 1964. "The performance of multi-man monitoring teams." *Human Factors*, 6, 179-184.

[19] Kanekar, S. 1982. "Individual and group performance on an anagrams task." *Australian Journal of Psychology*. 34(3), 337-344.

[20] Stanislaw, H. 1995. "Effect of type of task and number of inspectors on performance of an industrial inspection-type task." *Human Factors*, 37, 182-192.

[21] Baron, R. A. and D. Byrne (editors) 1997. *Social Psychology*. Boston: Allyn and Bacon.

[22] Langer, E. J. 1975. "The illusion of control." *Journal of Personality & Social Psychology*. 32, 311-328.

[23] DOE (U.S. Department of Energy) 2007. *General Corrosion and Localized Corrosion of the Waste Package Outer Barrier*. ANL-EBS-MD-000003 REV 03. Office of Civilian Radioactive Waste Management, Las Vegas, Nevada, July, 2007.

[24] INPO (Institute of Nuclear Power Operations) 2004. *Principles for a Strong Nuclear Safety Culture*. INPO, Atlanta, GA. November, 2004.

[25] Idaho National Laboratory 2011. *Systems Analysis Programs for Hands-on Integrated Reliability Evaluations (SAPHIRE) Version 8, Volume 1:Overview and Summary.* NUREG/CR-7039. Idaho Falls, Idaho June 2011.

Appendix A: Additional Analysis Details

Table A-1 List of all basic events in the 0705EARLYEND 2009 file and their probabilities – Yucca Mountain Analysis:

Name	Prob.	Description
INIT	1.00E+00	Placeholder initiating event
WP-LPB_CHECK	1.60E-01	Inspection fails to detect LPB failure after LPB process
HT_INSPECT_WP	1.60E-01	Inspection detects improper heat treatment cooldown?
WP_SHELL_TC_CHECK	1.60E-01	Checker fails to detect waste package shell thermocouple install error
WP_LID_TC_CHECK	1.60E-01	Checker fails to detect waste package lid thermocouple install error
WP-LPB_ACR	8.10E-02	Checker detects operator's failure to respond to annunciator
WP_SHELL_TC_INSTALL	8.10E-02	Waste package shell thermocouple improperly installed
WP_LID_TC_INSTALL	8.10E-02	Waste package lid thermocouple improperly installed
CRANE_MALFUNCTION	3.75E-03	Crane fails to move waste package from furnace to quench tank in specified time period
CRANE_OPERATOR_ERROR	3.75E-03	Crane operator error fails to move waste package from furnace to quench tank in specified time period
TROLLEY_MALFUNCTION	3.75E-03	Trolley fails to move waste package lid from furnace to quench chamber in specified time period
TROLLEY_OPERATOR_ERROR	3.75E-03	Trolley operator error fails to move waste package lid from furnace to quench chamber in specified time period
HT_OPERATOR_ERROR	3.00E-03	Heat treatment process operator fails to respond to alarm
WP-LPB-OPERATOR	3.00E-03	Operator fails to responds to annunciator
HT_LID_QUENCH_WP	3.00E-03	WP lid quenched OK?
MAKEUP_WATER_SYSTEM_FAIL	3.00E-03	Quench tank intake valve fails to open
CAM_DET	2.50E-03	Detection of unconformity on camera?
CHECK_BM_FLAW	1.25E-03	Inspection fails to detect base metal flaw
WELD_FILLER_ISP-WP	1.25E-03	Weld filler material inspection failure
SNORKEL_ATTACHMENT_FAIL	5.00E-04	Snorkel improperly attached to waste package
WP-LPB-IC	3.00E-04	Instrumentation & control system fails to alarm
TIMER_FAILURE	1.20E-04	Timer alarm fails to alarm
BM_FLAW	1.00E-04	Base Metal Flaw
WELD_FILLER_FLAW-WP	1.00E-04	Weld filler material flaw-Waste Package
MISH-BE-1	4.80E-05	Waste package mishandling 1
MISH-BE-2	4.80E-05	Waste package mishandling 2
MISH-BE-3	4.80E-05	Waste package mishandling 3
MISH-BE-4	4.80E-05	Waste package mishandling 4
MISH-BE-5	4.80E-05	Waste package mishandling 5
MISH-BE-6	4.80E-05	Waste package mishandling 6
MISH-BE-7	4.80E-05	Waste package mishandling 7
MISH-BE-8	4.80E-05	Waste package mishandling 8
WP-LPB-SENSOR	1.00E-05	Pressure monitor to LPB hydraulic system fails

Table A-2 Importance measure results for all early failure basic events for the end state DAMAGED-WP (sorted by uncertainty)

Name	#	Prob.	FV	RIR	RRR	Birn.	RII	RRI	Uncert.	Description
HT_OPERATOR_ERROR	7	3.00E-03	4.48E-01	149.	1.81	1.59E-02	1.59E-02	4.81E-05	1.18E-04	Heat treatment process operator fails to respond to alarm
WP-LPB_CHECK	3	1.60E-01	3.99E-01	3.10	1.66	2.68E-04	2.25E-04	4.29E-05	5.43E-05	Inspection fails to detect LPB failure after LPB process
WP-LPB_ACR	3	8.10E-02	3.99E-01	5.53	1.66	5.30E-04	4.87E-04	4.29E-05	5.43E-05	Checker detects operator's failure to respond to annunciator
WP-LPB-OPERATOR	1	3.00E-03	3.62E-01	121.	1.57	1.30E-02	1.29E-02	3.89E-05	9.60E-05	Operator fails to responds to annunciator
HT_INSPECT_WP	6	1.60E-01	1.54E-01	1.81	1.18	1.04E-04	8.71E-05	1.66E-05	2.10E-05	Inspection detects improper heat treatment cooldown?
CRANE_MALFUNCTION	2	3.75E-03	1.09E-01	29.9	1.12	3.12E-03	3.11E-03	1.17E-05	8.77E-06	Crane fails to move waste package from furnace to quench tank in specified time period
CRANE_OPERATOR_ERROR	2	3.75E-03	1.09E-01	29.9	1.12	3.12E-03	3.11E-03	1.17E-05	8.77E-06	Crane operator error fails to move waste package from furnace to quench tank in specified time period
TROLLEY_MALFUNCTION	2	3.75E-03	1.09E-01	29.9	1.12	3.12E-03	3.11E-03	1.17E-05	8.77E-06	Trolley fails to move waste package lid from furnace to quench chamber in specified time period
TROLLEY_OPERATOR_ERROR	2	3.75E-03	1.09E-01	29.9	1.12	3.12E-03	3.11E-03	1.17E-05	8.77E-06	Trolley operator error fails to move waste package lid from furnace to quench chamber in specified time period
HT_LID_QUENCH_WP	2	3.00E-03	7.13E-02	24.7	1.08	2.55E-03	2.55E-03	7.66E-06	9.70E-06	WP lid quenched OK?
MAKEUP_WATER_SYSTEM_FAIL	2	3.00E-03	7.13E-02	24.7	1.08	2.55E-03	2.55E-03	7.66E-06	9.70E-06	Quench tank intake valve fails to open
WP_SHELL_TC_CHECK	2	1.60E-01	6.75E-02	1.35	1.07	4.54E-05	3.81E-05	7.26E-06	9.19E-06	Waste package shell thermocouple checker fails to detect installer's error
WP_SHELL_TC_INSTALL	2	8.10E-02	6.75E-02	1.77	1.07	8.96E-05	8.23E-05	7.26E-06	9.19E-06	Waste package shell thermocouple improperly installed
WP_LID_TC_CHECK	1	1.60E-01	5.79E-02	1.30	1.06	3.89E-05	3.27E-05	6.22E-06	7.88E-06	Waste package lid thermocouple checker fails to detect installer's error
WP_LID_TC_INSTALL	1	8.10E-02	5.79E-02	1.66	1.06	7.68E-05	7.06E-05	6.22E-06	7.88E-06	Waste package lid thermocouple improperly installed
WP-LPB-IC	1	3.00E-04	3.62E-02	122.	1.04	1.30E-02	1.30E-02	3.89E-06	9.60E-06	Instrumentation & control system fails to alarm
TIMER_FAILURE	4	1.20E-04	1.67E-02	140.	1.02	1.49E-02	1.49E-02	1.80E-06	4.42E-06	Timer alarm fails to alarm
SNORKEL_ATTACHMENT_FAIL	2	5.00E-04	1.19E-02	24.7	1.01	2.55E-03	2.55E-03	1.28E-06	3.15E-06	Snorkel improperly attached to waste package

Name	#	Prob.	FV	RIR	RRR	Birn.	RII	RRI	Uncert.	Description
CAM_DET	8	2.50E-03	8.93E-03	4.56	1.01	3.84E-04	3.83E-04	9.60E-07	7.20E-07	Detection of unconformity on camera?
CHECK_BM_FLAW	1	1.25E-03	1.16E-03	1.93	1.00	1.00E-04	9.99E-05	1.25E-07	9.37E-08	Inspection fails to detect base metal flaw
WELD_FILLER_ISP-WP	1	1.25E-03	1.16E-03	1.93	1.00	1.00E-04	9.99E-05	1.25E-07	9.37E-08	Weld filler material inspection failure- Waste Package
BM_FLAW	1	1.00E-04	1.16E-03	12.6	1.00	1.25E-03	1.25E-03	1.25E-07	9.37E-08	Base Metal Flaw
WELD_FILLER_FLAW-WP	1	1.00E-04	1.16E-03	12.6	1.00	1.25E-03	1.25E-03	1.25E-07	9.37E-08	Weld filler material flaw-Waste Package
MISH-BE-1	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 1
MISH-BE-2	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 2
MISH-BE-3	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 3
MISH-BE-4	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 4
MISH-BE-5	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 5
MISH-BE-6	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 6
MISH-BE-7	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 7
MISH-BE-8	1	4.80E-05	1.12E-03	24.3	1.00	2.50E-03	2.50E-03	1.20E-07	2.96E-07	Waste package mishandling 8
WP-LPB-SENSOR	1	1.00E-05	1.21E-03	122.	1.00	1.30E-02	1.30E-02	1.30E-07	9.72E-08	Pressure monitor to LPB hydraulic system fails

Fussell Vesely (FV) is the fractional contribution of the basic event to the total risk: $FV_i = F_i(x) / F(x)$

Risk Reduction Ratio (RRR) or Interval (RRI) is how much risk would decrease if the basic event probability=0. RRI is also called inspection importance: $RRR = F(x)/F(0)$ -or- $RRI = F(x) - F(0)$

Risk Increase Ratio (RIR) or interval (RII) is how much risk would increase if the basic event probability=1.

$$RIR = F(1)/F(x) \text{ -or- } RII = F(1) - F(x)$$

Birnbaum is how much the total risk changes with respect to changes in basic event probability: $B = F(1) - F(0)$

Where $F(x)$ is the original probability of the end state (P(Damaged-WP)); $F_i(x)$ is the probability of the end state with only the basic event of interest; F(0) is the probability of the end state with the event probability set equal to 0 (perfectly reliable); F(1) is the probability of the end state with the event probability set equal to 1 (failed)

Appendix B: Detailed Recommendation for Task Analysis

This appendix (Table B-1) provides an example of the type of information that would be collected and analyzed through a detailed task analysis of a disposal overpack manufacturing process. Specifically, this table identifies information that is known, unknown, and that which should be considered to mitigate the "operator fails to respond to alarm" error during the "heat treatment of outer corrosion barrier" task.

The main goal of the task analysis is to determine the following information:

- Who is participating in the task or procedure?
- What information needs to be presented?
    - How should it be presented?
    - Who is it being presented to?
- What are the alternative courses of action given anomalies to the standard procedure?

An analysis of this information can provide designers with concrete information that can help drive the development of the given task or procedure, ensuring that human factors that can lead to (or mitigate) errors have been considered.

Table B-1. Representational of detailed task analysis for the disposal overpack manufacturing

Task: Heat Treatment of Outer Corrosion Barrier **Error**: Operator fails to respond to alarm		
Step 1: Place outer corrosion barrier in heat		
Known	**Unknown**	**Considerations**
Temperature: 2050F +/- 50F *Thermocouples used to measure temperature*	How is temperature displayed?	Design of display including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multi-modal displays, other tasks occurring simultaneously (operator workload), F vs. C, changeable between F vs. C
	Who is temperature displayed to?	Crane operator, supervisor, others?
	Is there an indication (alarm) that temperature is out of range?	Consider alarm design including size, color, brightness, viewing angles, environmental impacts (sun, dark), use of multimodal displays, other tasks occurring simultaneously (operator workload), sound, flash, rate
Time: 20 minutes minimum	How (mechanical) and in what units is time measured?	Timer Format of time indications (h:mm; hhmm; hh:mm; hhmmss; hh:mm:ss; etc…)
	How is time displayed?	Design of display including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multi-modal displays, other tasks occurring simultaneously (operator workload)
	Who is time displayed to?	Crane operator, supervisor, others?
	What is the indication to start the time? How does this occur? Is there an operator action associated with starting the time?	
	What are possible alternative (alarm) scenarios? What are the appropriate courses of action given these scenarios? Are there inappropriate actions that could aggravate this event? Is course of action information displayed, if so, how? Who is this course of action information displayed to if it is displayed?	
Step 2: Remove outer corrosion barrier from heat		
Known	**Unknown**	**Considerations**
Thermocouple temperature range shall be 2050F+/-50F for 20 minutes minimum	How is temperature displayed?	Design of display including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multi-modal displays, other tasks occurring simultaneously (operator workload), F vs. C, changeable between F vs. C
	Who is temperature displayed to?	Crane operator, supervisor, others?
	What is the indication that (time) 20 minutes is up and within the proper temperature range?	Design of display including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multi-modal displays, other tasks occurring simultaneously (operator workload); format of time indications (h:mm; hhmm; hh:mm; hhmmss; hh:mm:ss; etc…)
	What is the indication (alarm) that temperature is out of range?	Consider alarm design including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multimodal displays, other tasks occurring simultaneously (operator workload), sound, flash, rate, length of time, silence/mute, dismiss

	What are possible alternative (alarm) scenarios? What are the appropriate courses of action given these scenarios? Are there inappropriate actions that could aggravate this event? Is course of action information displayed, if so, how? Who is this course of action information displayed to if it is displayed?	
Step 3: Transition outer corrosion barrier from heat to cooling		
Known	**Unknown**	**Considerations**
Temperature must remain above 2020F during transition	How is temperature displayed?	Design of display including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multi-modal displays, other tasks occurring simultaneously (operator workload), F vs. C, changeable between F vs. C
	Who is temperature displayed to?	Crane operator, supervisor, others?
	What is the (alarm) indication that temperature is out of range?	Consider alarm design including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multimodal displays, other tasks occurring simultaneously (operator workload), sound, flash, rate, length of time, silence/mute, dismiss
	What are possible alternative (alarm) scenarios? What are the appropriate courses of action given these scenarios? Are there inappropriate actions that could aggravate this event? Is course of action information displayed, if so, how? Who is this course of action information displayed to if it is displayed?	
Step 4: Place outer corrosion barrier in cooling process		
Known	**Unknown**	**Considerations**
Corrosion barrier temperature must be 2020F or higher when entering cooling	How is temperature displayed?	Design of display including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multi-modal displays, other tasks occurring simultaneously (operator workload), F vs. C, changeable between F vs. C
	Who is temperature displayed to?	Crane operator, supervisor, others?
	What is the (alarm) indication that temperature is out of range?	Consider alarm design including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multimodal displays, other tasks occurring simultaneously (operator workload), sound, flash, rate, length of time, silence/mute, dismiss
	What are possible alternative (alarm) scenarios? What are the appropriate courses of action given these scenarios? Are there inappropriate actions that could aggravate this event? Is course of action information displayed, if so, how? Who is this course of action information displayed to if it is displayed?	

Cooling rate must be greater than 275F/min	How is time and temperature (rate) displayed?	Design of display including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multi-modal displays, other tasks occurring simultaneously (operator workload); format of time indications (h:mm; hhmm; hh:mm; hhmmss; hh:mm:ss; etc…), F vs. C, changeable between F vs. C
	Who is time and temperature (rate) displayed to?	Crane operator, supervisor, others?
	What is the (alarm) indication that the cooling rate is insufficient?	Consider alarm design including size, color, brightness, viewing angles, environmental impacts (sun, dark, vibration), use of multimodal displays, other tasks occurring simultaneously (operator workload), sound, flash, rate, length of time, silence/mute, dismiss
	What is the process if the cooling rate falls below 275F/min?	
	What are possible alternative (alarm) scenarios? What are the appropriate courses of action given these scenarios? Are there inappropriate actions that could aggravate this event? Is course of action information displayed, if so, how? Who is this course of action information displayed to if it is displayed?	

Appendix C: HRA Data Collection

There are two fundamental HRA data issues. The first is the lack of relevant HRA data. A large amount of human performance data exists. This data ranges from low-level, specific performance information of single elementary tasks to high-level, generic knowledge of human behavior in both qualitative and quantitative forms. The data includes both empirical, raw data from facilities and industries and derived - or estimated - data obtained in statistical or HRA analyses with or without support of expert judgment. However, there is still a chasm between the available human performance data and HRA data needs. That is, while the current human performance data can be used to inform HRA, more data are needed for HRA to predict possibilities of operation failures, operator response opportunities to operation failures, plant sensitivity to operator response, and various factors that drive human reliability under the extreme conditions considered in PRA. Currently, the various forms of human reliability data are primarily used as a basis for expert judgment rather than as an input to HRA applications. This situation is partly because much of the data are often collected to improve operator performance without a focus on HRA, and partly because existing HRA databases, developed in isolation, focus on different aspects of HRA data. This calls for the need for a definition for HRA data (i.e. what constitute HRA data). All in all, HRA data exploitation is a discovery science. What we discover comes back to the question that we ask ourselves when we start the exploitation – what do we want the data to provide? The implication is that we need to understand what data we need for waste package manufacturing and handling. Such an understanding will serve as the basis for developing a data collection framework. The importance of understanding the waste package manufacturing and handling process is implicated by the fact that the requirements on data types and the level of information detail are largely determined by the intended applications of the data. Since data are needed for HRA method implementation and/or HRA/PRA review for DCP early failure probability quantification, data collection effort may need to focus on scenario- or task-specific data.

The second HRA data issue is the lack of ability to leverage data collected on one HRA theoretical perspective for use by methods that are based on a different theoretical perspective. This is partly because of the lack of a common theoretical basis for data collection, analysis, and exchange. Although HRA methods agree on many of the central tenets necessary for conducting an HRA and share similarities in the systems, operations and accident conditions to which they are applied, they differ in their theoretical bases, underlying human performance models, and analysis approaches. The method diversity implies that, while the methods share common data needs, distinctions in the respective HRA quantification models place method-specific data requirements to support method implementation and development. One implication for waste package manufacturing and handling is that the data needs and the data collection framework will, to some degree, be dependent on the HRA method we decide to use. Regardless if we decide to develop a new HRA method or adapt an existing method, it is important to strengthen the theoretical basis and techniques for data exploitation to integrate data and theory and yield new insights.

There are diverse data taxonomies, terminologies, or categorizations used in HRA databases to acquire data. Each taxonomy has an emphasis on data of different types and at different information detail levels. Information that is beyond the scope of a taxonomy is normally not

recorded in the respective database, even though the information is considered relevant or important from the perspective of another database. For HRA data to have a broader influence, a generic, neutral data taxonomy is needed as an overarching theoretical context to organize extant data, guide future data collection activities, and serve as a common language for data exchange. Such a taxonomy should also be an open framework for inclusion of new information to accommodate the broad scope of HRA data and the diverse specificities of data applications. In addition, the generic, neutral data taxonomy should treat tasks at a micro-cognitive and generic failure mode level rather than at a micro-cognitive and specific (e.g., turning a control knob) level.

Since operators are more likely to fail in cognitively challenging situations, future HRA data collection needs to be more focused on data concerning cognitive failures and associated contextual factors that account for the development of inappropriate plant status assessment, goals, action plans, and the tendency to persist in their implementation even when contradicted by observation. Data are also needed regarding cognitive demands and challenges in unexpected situations and potential crew response to confusing conditions.