Guidelines for post-intervention analyses of the original WHI trials

WHI offers a unique opportunity to evaluate longer term effects of clinical trial randomization to low fat diet and/or hormone therapy and/or calcium and vitamin D supplementation on health outcomes. The fact that the trial interventions ended in 2005 or earlier, participants who were not reconsented were lost to follow-up and changes were made in the follow-up protocol, and the fact that these data have been examined previously makes it necessary to establish guidance for analytical plans and final manuscripts to assure the highest quality science is published.

The following factors should be addressed in a robust analysis plan of longer term or late effects of the original WHI interventions:

1. Cohort attrition over time
2. Changes in WHI outcomes ascertainment and adjudication and associated measurement error issues. The protocol was streamlined to r
3. Use of NDI versus proxy report of cause of death (note: fall 2014 currently available, complete NDI dataset)
4. Adherence to and cross-over between arms and potential differences in behaviors or exposures related to the intervention (e.g., differential health screening, uptake of other medications) during the post-intervention phase. As well as lack of data regarding use/adherence to HT, diet and/or dietary supplementation (calcium, vitamin D) post-trial end
5. Multiple testing by virtue of prior analyses of the outcome by examining a new outcome not previously specified
6. Power, including likelihood of an intervention effect persisting 10 or more years after intervention cessation
7. Differential follow up for incident disease
8. Interpretation, given factors #1 to #7, and given intervention effects over the preceding post-randomization time period for analyses having a follow-up period that begins many years after randomization.

<u>Guidance for addressing these factors for analyses using post-intervention data:</u>

<u>Cohort Attrition</u>: Participants may have been lost to follow-up for multiple reasons. Analyses relying on active follow-up after trial termination should examine the potential for differential follow-up and thus outcomes ascertainment bias between arms. Statistical methods to address cohort attrition should at a minimum include time-dependent stratification for study phase. Risk factors related to the outcome of interest and any data available on intervention adherence, cross-over, or other exposures should be examined by randomization assignment among those remaining in the cohort at each key follow-up point (beginning of post-intervention, 2005, 2010, 2015). More sophisticated analyses, including inverse probability weighted analyses, may be needed if noteworthy imbalances are found. For mortality outcomes that can rely on linkage to NDI, which is available for nearly all participants through linkage regardless of re-consent status, these concerns about cohort attrition do not apply. However, mortality results should end with an NDI linkage to assure nearly complete ascertainment of vital status through a specific date.

<u>Changes in outcomes procedures</u>: The protocols for most outcomes data collection other than cancer have changed over time. Table 1 below, from the Data Preparation documentation on WHI.org shows the highest quality data available for each outcome by study period. While other data sources are available (Self or proxy

report, Medicare and NDI) these are considered to be of lesser quality because of the less rigorous review, more limited supporting documentation and the lack of details on the diagnosis. If analyses will use outcomes from multiple sources information, investigators should examine the correspondence of events among those participants who have outcomes data from multiple sources (e.g. adjudicated and Medicare). In some cases, such analyses may already be published and references to these and an acknowledgement of the likely impact of outcome ascertainment errors on the current analysis may be all that is needed. Efforts to incorporate validation study results into the current analysis would be ideal.

Table 1: WHI adjudicated outcomes by study period and WHI cohort

	WHI 1993-2005				Extension 1 2005-2010	Extension 2 2010-2015	Extension 2 2010-2015	
Outcome	HT	DM	CaD	OS	CT/OS	MRC	SRC	**Form**
CARDIOVASCULAR:								
MI	C	L	L	L	C	C	S	121
Stroke	C	L[1]	L[1]	C[1]	C	C	S	121/132
WHI congestive heart failure	C	L	L	L	S	S	S	121
UNC heart failure	C[2]	C[2]	C[2]	C[2]	C[2]	C[2]	S	135/136
Angina	C	L	L	L	S	S	S	121
Peripheral artery disease	L	L	L	L	C	C	S	121
Carotid artery disease	L	L	L	L	C	C	S	121/132
Coronary revascularization	L	L	L	L	C	C	S	121
TIA	C	L	L	L	S	S	S	121/132
Atrial fibrillation						C	S	121
Heart valve disease						C	S	121
Aortic aneurysm/dissection						C	S	121
CANCER:								
Breast cancer	C	C	C	C	C	C[3]	C[3]	122/130
Endometrial cancer	C	C	C	C	C	C[3]	C[3]	122/130
Colorectal cancer	C	C	C	C	C	C[3]	C[3]	122/130
Ovarian cancer	C	C	C	C	C	C[3]	C[3]	122/130
All other cancers	C	C	C	C	C	C[3]	C[3]	122/130
FRACTURES:								
Hip	C	C	C	C	C	C	S	123
Non-hip fractures	L	L	L	L[4]	S	S	S	123
OTHER:								
Pulmonary embolism	C	S	S	S	C (HT)	C	S	126
Deep vein thrombosis	C	S	S	S	C (HT)	C	S	126
Hysterectomy	C	S	S	S	C (HT)	S	S	131
Death from any cause	C	C	C	L	C	C	S	124

C - Centrally adjudicated
L - Locally adjudicated
S - Self-reported

[1] Used central adjudication data when available, and local adjudication data otherwise
[2] Adjudication by UNC is only done for the HT and Black/Hispanic participants. Data for events in Extension 2 is not yet available.
[3] All cancers centrally adjudicated on all Extension Study 2 participants
[4] Done only at the BMD centers

Interventions: Investigators should examine and report on adherence/use of the interventions, behavior changes or use of alternative therapies during the entire time interval during which outcomes are being examined to the extent possible (e.g., DM subsample with follow-up repeat 24 hour recalls; reported HT use

after trial).. Sensitivity analyses that use this information may be warranted if there is considerable cross-over or differential uptake of other approaches. ITT intervention period and post-intervention period effects should be examined separately. For blinded trials, reports should make it clear that participants were unblinded at the end of the intervention phase.

Multiple testing: No multiple testing adjustments have been developed to address the concerns regarding reporting of subsequent results after the primary analyses have been reported. To limit potential type I error inflation, a structural approach is recommended. Analyses should be based on a designated post-trial database with most recent generally applied unless the authors can provide a clear rationale for use of an earlier database: (Events through 2005, 2010, 2015, and if appropriate, subsequent 5 year increments until follow-up ceases).

Manuscripts should describe the results of any prior analyses for that outcome (with emphasis on the primary ITT analysis), including the fraction of current information (number of events) included in prior analyses and the duration of the intervention and post-intervention period. In general these should be characterized as descriptive analyses of longer-term effects in the absence of continuing intervention rather than hypothesis testing and should provide effect estimates for both the intervention and post-intervention periods. For new outcomes not defined in the trial protocol, the results should be clearly described as hypothesis generating. Other cautions related to subgroup analyses and intermediate outcomes still apply.

Power: Authors should discuss limitations of statistical power for testing interventions in the post-intervention and cumulative periods, which may vary by outcome/biologic mechanism. When using data from the post-intervention period, the interpretation these analyses should be informed by the factors noted above should be explicitly stated. Power limitations may be particularly important for mortality outcomes. Mortality outcome comparisons among randomized groups mostly lacked power during the intervention period, and can also be expected to lack power post-intervention since any intervention mortality effects may dissipate within a few years following cessation of intervention activities.

Interpretation: Comparisons among randomized groups that include post-intervention follow-up should comment on each of the above areas and, additionally, need to discuss the implications of selection deriving from treatment effects in the intervention period for analyses based on follow-up times that initiate at a post-intervention time. In general, these analyses should be characterized as providing definitive ITT comparison information only if outcomes over the entire follow-up period are presented, and each of the above issues is adequately addressed.

Procedures:

All manuscripts proposed must include a checklist (enclosed) that demonstrates the proposal or manuscript has addressed all issues listed along with the proposal/manuscript submitted to WHI P&P Committee.

All manuscripts/proposals submitted that include extended follow-up data and address mortality (all cause, total, disease-specific) must use the dataset that includes NDI data for cause of death through the most recent linkage. Manuscripts already submitted to a journal for review are exempt unless reviewers comment on the need for updated data in which case updated datasets should be used for analysis.

Checklist for analyses of CT results using post-intervention data

Item	Completed Y = Yes N= No NA = Not Applicable
Description of cohort attrition over time	Y (page ____) N NA
Assessment of potential differential follow-up by randomization assignment.	Y (page ____) N NA
Analysis uses time-dependent stratification by study phase (active intervention, post-intervention through2010, after 2010)	Y (page ____) N NA
Effect estimates for intervention and post-intervention periods as well as cumulative results provided and discussed	Y (page ____) N NA
Modifiable risk factors for outcomes of interest examined for differences between arms over follow up (e.g., differential uptake of medication, use of screening)	Y (page ____) N NA
Consideration of inverse probability weighted analysis when imbalances identified	Y (page ____) N NA
Use of NDI for mortality analyses	Y (page ____) N NA
Use of adjudicated outcomes whenever possible, description of source of outcomes	Y (page ____) N NA
Incorporation of “validation” results when using non-adjudicated outcomes	Y (page ____) N NA
Report of adherence to trial intervention during and post-trial (e.g. DM -subset with follow-up 24 hour recalls post-intervention). Sensitivity analysis if cross-over or differential uptake identified	Y (page ____) N NA
Clarification that participants were unblinded to intervention as of 2005	Y (page ____) N NA
Analysis is based on a previously agreed upon database (2010, 2015, and 2020). Most complete dataset at the time of each follow-up period should be used.	Y (page ____) N NA
Interpretation acknowledges multiple testing over time	Y (page ____) N NA
Interpretation addresses statistical power issues as appropriate to time interval examined (e.g., 10 years after intervention has ceased)	Y (page ____) N NA