Preliminary exploration of potential paths of dose and psychological response

Robert Alan Yaffee
Silver School of Social Work
New York University

robert.yaffee@nyu.edu

30 March 2012

Outline 2nd edition

1. Introduction
2. Trends in Risk perception
3. Dose-psychological response relationships
4. Reconstructed dose and medical diagnosis

Associated computer files: This file: RiskPerceptionDoseResponse.pdf
data: ch3wMaster27mar2012.dta
dofiles: TrendsInRiskPerception= wavegraphs.do and GraphingRiskTrends.do
Dose-Psychological Response relationships = doseresponse.do
= XtDoseResponse.do
Reconstructed dose and medical diagnosis = PanelRegIcdx.do
Dose-Psychlogical Response CS_PTSD= prelimXtscan3.do

Acknowledgments
We would like to thank Mariya Burdina for her assistance with data management for this project, as well as Remi Frazier and Professor Thomas Borak for their help in dose reconstruction of the depositions of $Cesium^{137}$. We would like to thank the National Science Foundation for making this research possible with HSD Grant #0826983.

1 Introduction

In this research note, we search for possible relationships between risk perception trend over time. We view risk perception from the point of view of the individual respondent, from his perspective of the threat to his family, and then to the extent that regional pollution stems from Chornobyl radiation.

We define our subject matter by examining the extent to which risk perception constitutes a global or a broken trend. A global trend is a systematic change in the mean over time, the slope of which is time invariant. A broken trend is similar to a spline, which constitutes one slope from wave one to wave two and then a different slope from wave two to wave three. In order to define the nature of a trend which entails a consistent, systematic change in the mean, we examine paneled lowess plots for males and females over our three waves of 1986, 1987 thru 1996, and 1997 through the time of the interview.

When we examine the dose-psychological response relationships, we do this in two sections. In the first section, we use a reconstructed dose, measured in microGrays (μGrays), by average cumulative dose per wave, whereas the psychological response is measured by a set of well-known psychological scales, the reliabilities we provide. We perform a Prais-Winston regression of the scales against the wave measure to test whether or not there is a significant trend in the scale.

If we find a significant trend in the scale, we will have to detrend the variable before performing an analysis on it. We may detrend the variable by first differencing it, lest we engender spurious regression. We perform separate analysis for males and females and will transform as necessary only to render the analysis feasible. In order to understand this concept, we have to come to an concurrence with regard to nomenclature. Does a trend include an intercept as well as a slope or is a trend limited to the construct of a slope? A broad construction of a trend would include both an intercept and a slope, whereas a strict construction might limit a trend to the construct of a slope alone. Because a first difference of a level is a slope, one can be easily converted to the other, and both might be a trend, depending upon whether one is analyzing levels or rates. At a minimum, a trend includes a slope, and it might also include an intercept under particular circumstances. We generally find the broad construction being used in the literature on state space models for which reason, we will adhere to that convention here.

We endeavor to examine candidate pathways of dose and psychological response, dose, mental health, and then risk-perception and dosage. Where dosage is measured in microGrays and where possible physical and mental health impacts entail psychological symptomatology as measured by the Nottingham health profile and the Basic Symptom inventory sub-scales, we endeavor to reveal possible pathways of the relationships through which these manifestations appear.

2 Trends in Risk perception

The subject of this paper is a study of the trends in risk perception on the part of the nuclear disaster survivors. To explain the trends, we first depict such changes in terms of Lowess plots of the risk perceptions over the three waves of

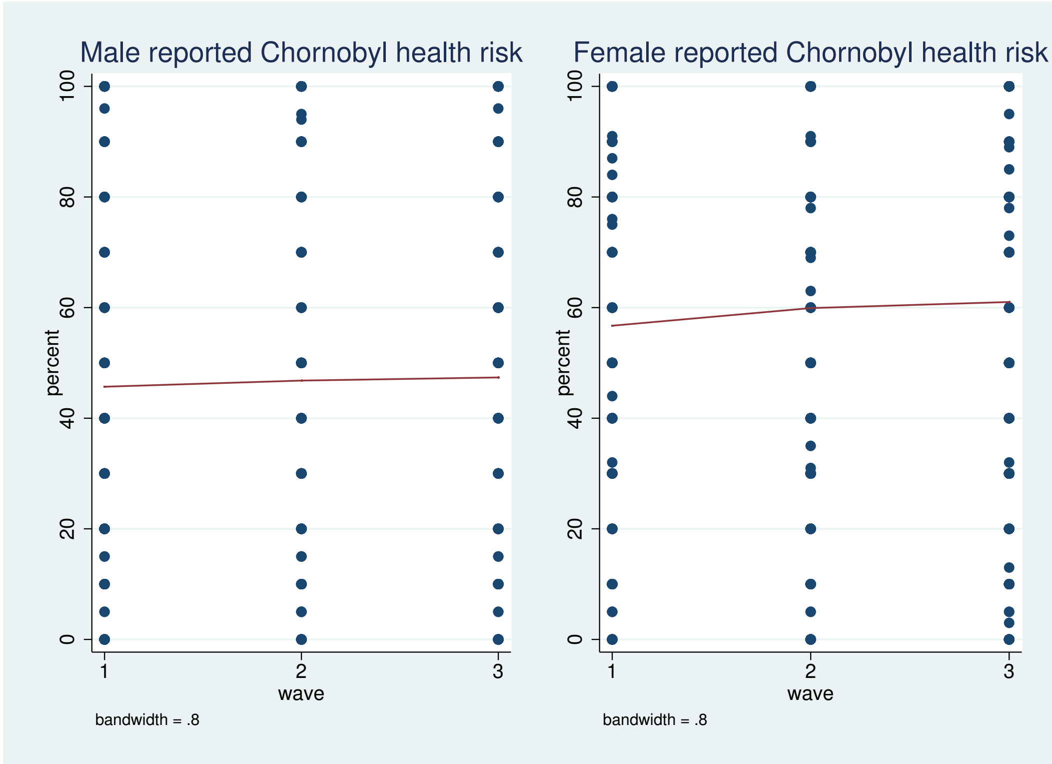

Figure 1: Male and Female reports of Chornobyl related health risk

our study and then we try to reveal these changes over time while displaying the size of their confidence intervals by which we can determine whether these changes constitute significant differences. Figure one displays a lowess plot of the Chornobyl related risk perception to oneself for male and female respondents.

Figure two shows the confidence intervals of the trends over our three waves. Both of these graphs reveal evidence of a broken or local trend. For the males, the level of self-perceived Chornobyl health risk increases from 45.7% to 46.8% to 47.4% over waves one, two, and three, respectively. In contrast, the female level increases from 56.7% to 59.9% to 61.01%, respectively, over these same three waves. To determine whether these confidence intervals reveal a significant trend, we not only offer Figure two, but we provide an analysis of a Prais-Winston regression, which corrects for autocorrelation bias in the data.

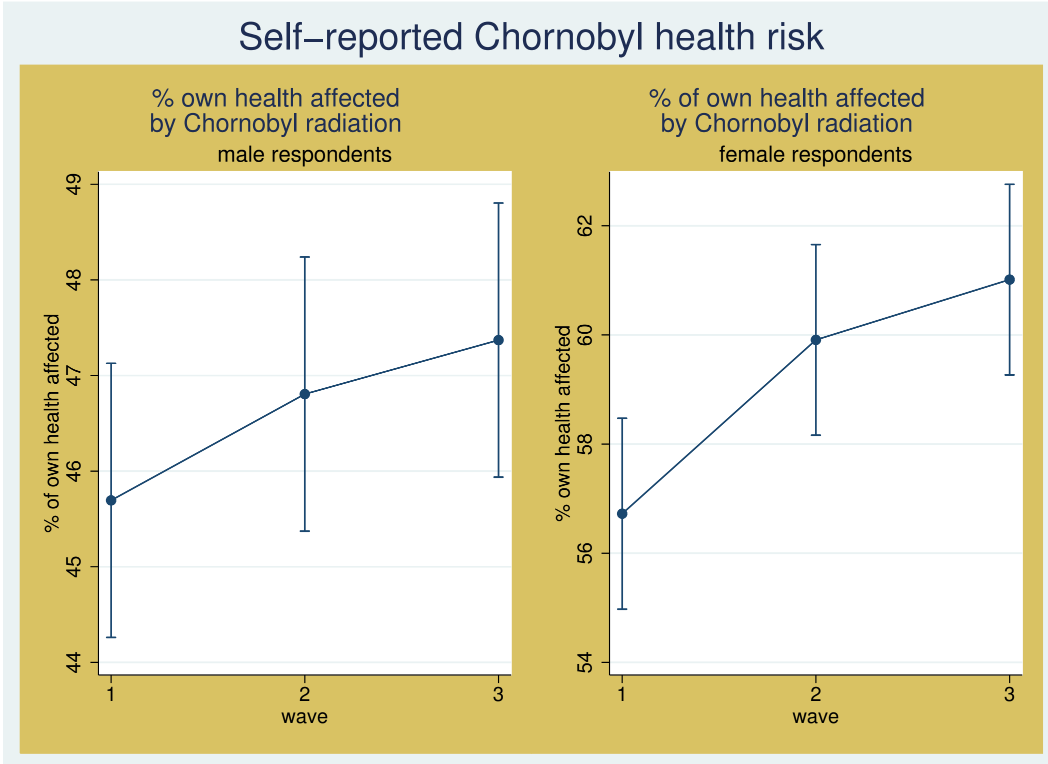

Figure 2: Male and Female reports of Chornobyl related health risk

Table 1 Male Regression of Self-Perceived Chornobyl health risk again wave

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	53686.911	1	53686.911
Residual	270514.931	1018	265.73176
Total	324201.842	1019	318.156862

Number of obs =	1020
F(1, 1018) =	202.03
Prob > F =	0.0000
R-squared =	0.1656
Adj R-squared =	0.1648
Root MSE =	16.301

radhlw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	.8382353	.6251256	1.34	0.180	-.3884468	2.064917
_cons	44.87161	2.140397	20.96	0.000	40.67152	49.07171
rho	.8779693					

Table 2 Female Regression of Self-Perceived Chornobyl health risk again wave

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	95888.5564	1	95888.5564
Residual	431980.989	1086	397.77255
Total	527869.546	1087	485.620557

Number of obs	= 1088
F(1, 1086)	= 241.06
Prob > F	= 0.0000
R-squared	= 0.1817
Adj R-squared	= 0.1809
Root MSE	= 19.944

radhlw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	2.210414	.7407105	2.98	0.003	.7570282	3.663799
_cons	54.48048	2.183304	24.95	0.000	50.19651	58.76446
rho	.8056074					

The implications are that prior to extensive modeling, both of these measures have to be mean centered, whereas the female version needs detrending as well. Unless we transform these variables accordingly, they may not be stationary and we will risk spurious regression bias as noted by Clive Granger and Paul Newbold in their book on Forecasting in Business and Economics[6, 130].

If we had more time periods, we could employ a state space model which using an augmented Kalman filter to estimate the model. However, to overcome the challenges of a nonstationary series by a sequential updating of a noninformative prior distribution with the data at each period until convergence to a steady state is reached generally requires more than three periods. A particle filter might be more appropriately used in this case, if there are not too many outliers. Particle filters use a noninformative prior density and a transition density to generate an empirical density, from which a sampling/importance resampling (developed by Rubin in 1987), along with rejection sampling (similar to that of the Metropolis algorithm) , and Markov Chain Monte Carlo estimation to fit non-Gaussian non-linear models [9].

2.0.1 a caveat on risk perception as a dependent variable

To avoid nonsensical regression effects, deriving from spurious correlations between independent and dependent variables, we need to take some precautions, while we are performing panel or even survival analysis. Many of the classical frequentist assumptions are based on the assumption that the distributions involved–particularly, of the model residuals are normally or multivariate normally distributed. The Gaussian or normal distribution is a two parameter distribution that can be defined by its mean and variance. For that reason, the mean and variance are sufficient statistics for defining a normal distribution. They are, in a manner of speaking, the instruments with which we will be doing some of our analysis.

To avoid unreliable or irreproducible results, the mean and variance of our analysis, regardless of the wave, must be stable over time. Similarly, the autocovariance must be stable over time. If these conditions hold, the processes with which we are dealing are covariance stationary. For conventional time series processes as well as time series cross-sectional analysis, covariance stationarity is a *sine qua non*.

The implication is that unless we engage in processes sequential importance sampling, or Gibbs sampling, or other Bayesian techniques to analyze non-Gaussian and nonlinear processes, we are constrained to adhere to conditions of covariance stationarity in analyzing series, unless we have feasible alternatives.

To avoid the trap of spurious regression, where both the dependent variable and the independent variable are functions of time, but otherwise are totally unrelated to one another, a regression of one upon the other might yield a high R^2 but a dubiously small Durbin-Watson d statistic, signifying a relationship of the first order autocorrelation. Under these circumstances, by dint of the common trend, the two variables appear to be related to one another, whereas they have no effect on one another at all in reality. In short, a controlled experiment would show that a change in one would not generate a change in the other, notwithstanding their spurious correlation. When Clive Granger and Paul Newbold revealed this relationship, they suggested that only by detrending both variables, could one analyze the effect of one upon the other without that relationship being confounded by spurious causality[9].

We may not be able to detrend our focal series by first differencing without losing all information about the source of information–namely, the Chornobyl disaster– which takes place in our first wave. We will therefore try to be sure that the processes we are dealing with do not entail random walks or other stochastic trends which may confound our analysis. Therefore, we may have to detrend by including a proxy measure for it as an independent variable in the model. Moreover, we will endeavor to show that any time series regression technique employed does not encounter a unit root problem and that we control for autocorrelation in our analysis whether it be a form of panel data or time series analysis.

If detrending means mean-centering the variable, we will reduce the chances of encountering multicollinearity with the constant or mean of the equation. However, in the process we will be divested of the location parameter. However, if detrending means that we have have to first difference, if the autocorrelation parameter does not approach that of a unit root, we may have to content ourselves with controlling for the fixed or deterministic effect of the trend, without addressing the stochastic nature of it, short of resorting to Bayesian analysis with a uniform prior distribution.

When testing the radhlw for the males and the radhlw as dependent variables in a panel data analysis, we find that they exhibit a unit root

2.0.2 Testing for significance of a trend

If we perform a Prais-Winston regression, which contains a first-order correction for the autocorrelation in the series, to minimize serial correlation bias that is likely to exist within our panel data, we find that the slope is not statistically significant, such a result would indicate a constant or stable mean. From Table one we can see that the male model has a stable mean but has a nonzero constant or mean, whereas Table two shows that the female model has both a significant positive slope and a significant non-zero constant.

When we endeavored to perform a panel unit root test using the Hadri test on the series to determine whether it needed first differencing, the results indicated that there might be some panels with unit roots in among the males, and that the three panels were insufficient in number for the test to make a determination for the women. If we were to first difference the series, we would lose our observation relating to the temporal origin of the problem–namely, 1986 when the disaster took place. Therefore, we attempt to control for the trend by including it in the model. While this will control for a deterministic trend, it will not control for a stochastic trend. Because these models are preliminary indicators, we will try to make do with this approximation, given the fact that the trend is not a steep one.

2.1 Reported Chornobyl-related familial health risk

We also examine the health risk to the family as reported by the respondent, from the male as well as the female points of view. This measure has the variable name of $radfmw$. Lowess plots of these relationships are shown in Figure Three and marginal plots with confidence intervals of the same relationships are displayed in Figure Four. Both sets of graphs reveal broken positive trends where the slope from wave one to wave two, is steeper than that from wave two to wave three. Although the slope is larger in the first interim period than it is during the second interim period, this is what is to be expected.

For the males, the mean level of % of family health affected by Chornobyl radiation goes from 50.9 % to 56.01% to 56.73% over the three waves respectively. The female mean level begins at 60.06% in wave one and rises to 67.9% in wave two, and then to 68.6% in wave three.

Although the steepness of the slope levels off somewhat during the period between waves two and three, the question arises as to whether it goes from being a significant increase to a nonsignificant increase.

We need an objective criterion according to which we can determine whether these constitute statistically significant trends. To be sure, both exhibit significant intercepts at wave one. Whether the first segments of the broken trends are statistically significant is an open question to which we have to refer to regression output for the answer.

In Tables three and four, we present the male and female Prais-Winston regression output, which may provide sufficient information with which we can answer that question.

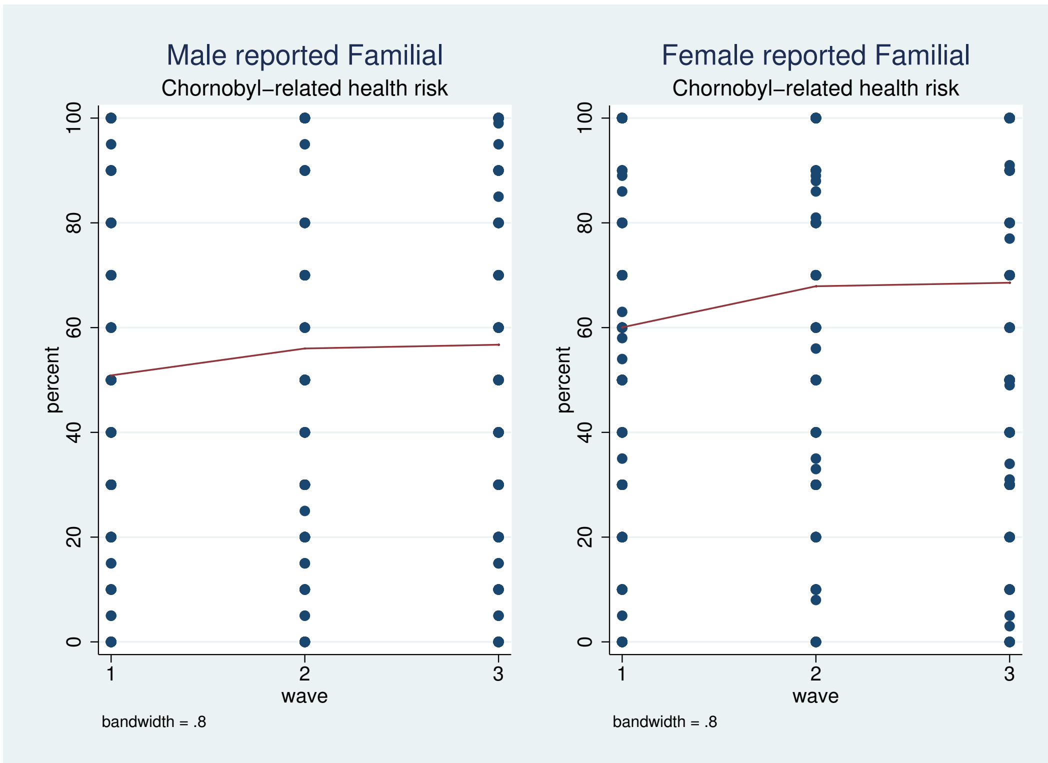

Figure 3: Male and Female reports of Chornobyl related family health risk

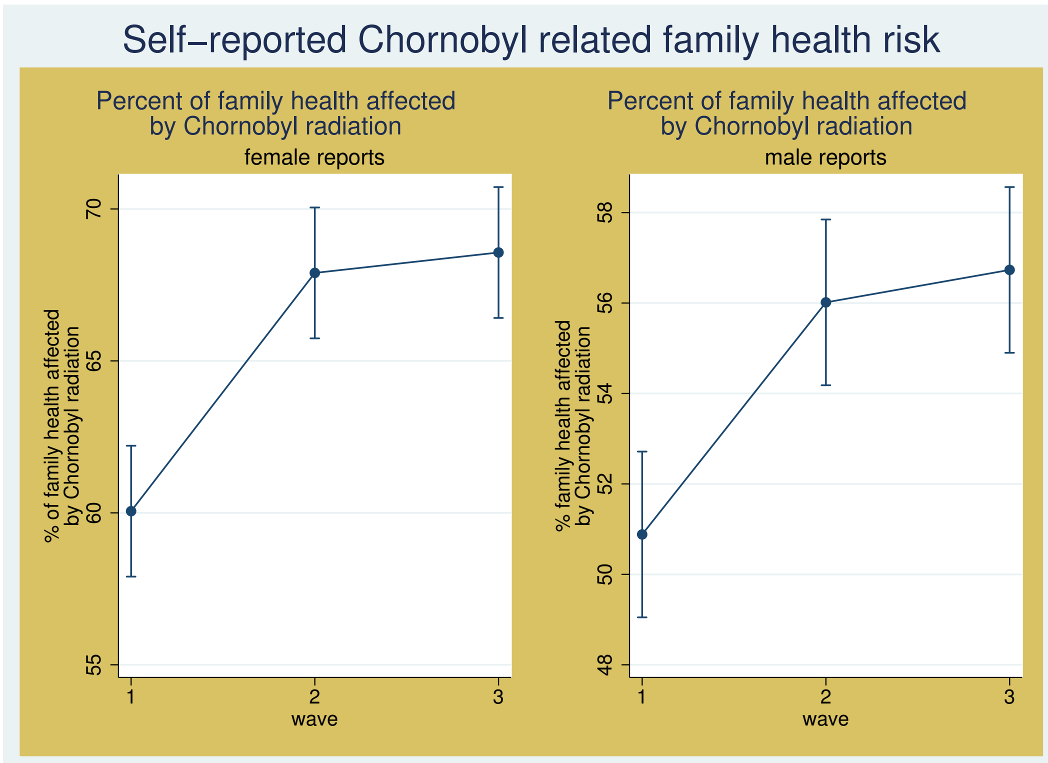

Figure 4: Male and Female reports of Chornobyl related family health risk

Table 3 Male Perception of Chornobyl health risk to the family

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	61891.4823	1	61891.4823
Residual	405644.004	1018	398.471517
Total	467535.486	1019	458.817945

Number of obs =	1020
F(1, 1018) =	155.32
Prob > F =	0.0000
R-squared =	0.1324
Adj R-squared =	0.1315
Root MSE =	19.962

radfmw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	2.925	.7654982	3.82	0.000	1.422865	4.427135
_cons	48.14155	2.30075	20.92	0.000	43.6268	52.65631
rho	.8179058					

Table 4 Female Perception of Chornobyl health risk to the family

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	88114.8736	1	88114.8736
Residual	634909.342	1087	584.093231
Total	723024.216	1088	664.544316

Number of obs =	1089
F(1, 1087) =	150.86
Prob > F =	0.0000
R-squared =	0.1219
Adj R-squared =	0.1211
Root MSE =	24.168

radfmw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	4.256198	.8969594	4.75	0.000	2.496231	6.016166
_cons	56.28039	2.337086	24.08	0.000	51.69468	60.8661
rho	.6896143					

From Tables 3 and 4, we see that both male and female trends (intercepts as well as slopes) are statistically significant, as indicated by the p-values for the wave variable being less than 0.05. It is easy to demean these measures, but we lose the location of the intercept if we do. If we first difference, we lose the point of primary reference. But if we use this measure as a dependent variable, we may have to first difference and speak about the rate rather than the level of this measure in order to be able to trust our analysis as being robust to spurious regression.

2.2 Cumulative exposure to radiation over one's lifetime

A possible third measure of risk is the self-assessment of lifetime cumulative exposure to radiation, as reported by men and women separately. We call this variable *radltw*. In Figure Five, the male and female perception of lifetime risk to themselves is plotted against the wave variable. What is very interesting in

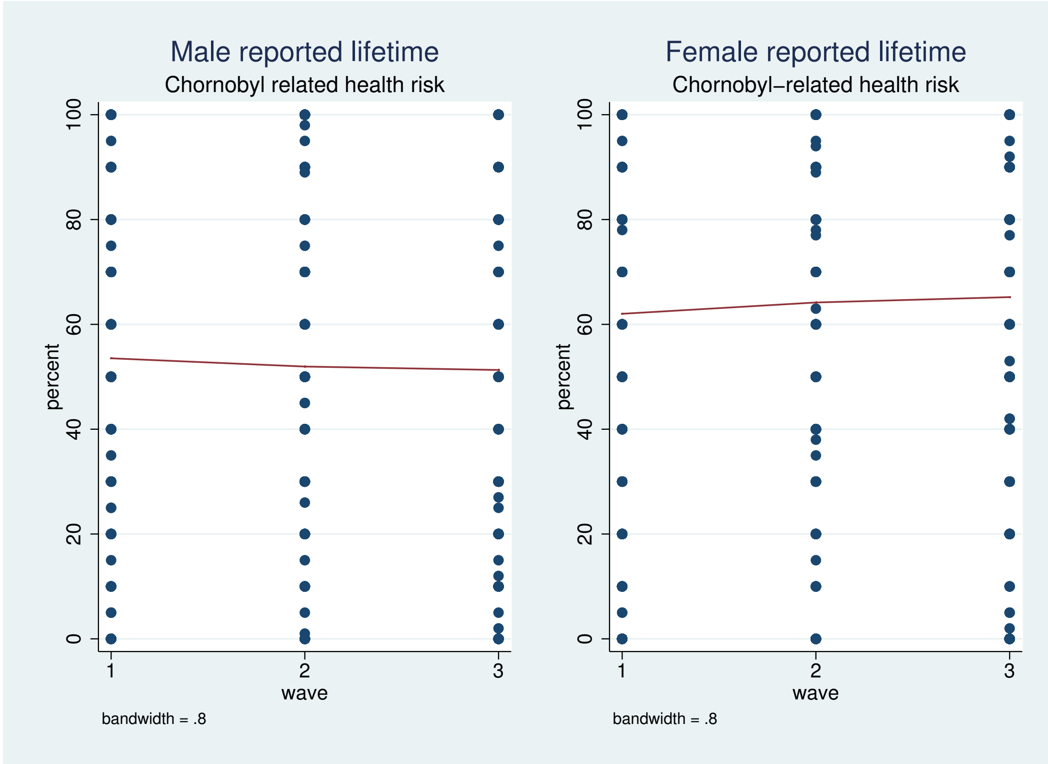

Figure 5: Male and Female reports health risk of lifetime exposure to radiation

this case is that according to the men, their risk seems to decline over time, whereas according to the women it seems to rise over time.

The mean levels of percent of health cumulatively exposed over one's lifetime differ for men and women. For the men, the level appears to decline over time as one proceeds from wave to wave. The males in wave one report 53.53%, in wave two they report 51.95%, whereas in wave three they report 51.3%. However, the women report what appears to be an increasing trend. In wave one, the females report 62.02%, and in the next wave they report 64.17%. By wave three, they report 65.2%. But are these trends statistically significant?

When we examine the graphs along with the data from Tables five and six, we observe that there is no statistically significant negative slope for the males because of the wide confidence intervals. The same cannot be said for the women. They exhibit a positive statistically significant slope (b = 1.59, p=0.027*) as well as a constant term that is statistically significantly different from zero ($\_cons = 60.48, p = 0.000^{***}$). In neither case is the autocorrelation coefficient high enough to be confused with a unit root: The male $\rho = 0.844$ and the female $\rho = 0.819$.

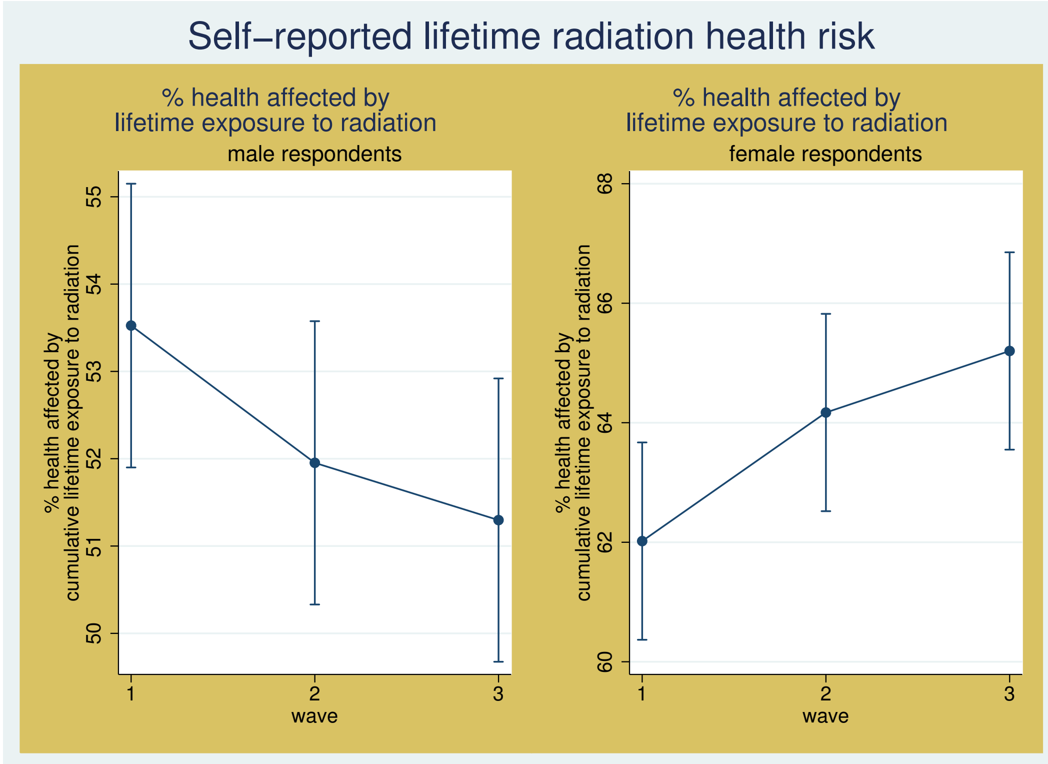

Figure 6: Male and Female reports of % health affected by lifetime cumulative exposure to radiation

Neither slope is steep, so the question arises whether they are statistically significant. It might be expected that the intercepts are statistically significant from zero, but can the same be said for their respective slopes? By examining Tables 5 and 6, we should be able to get an answer to this question.

Table 5 Male perception of health risk from cumulative exposure to radiation

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	105216.634	1	105216.634
Residual	345773.858	1017	339.993961
Total	450990.492	1018	443.016201

Number of obs	= 1019
F(1, 1017)	= 309.47
Prob > F	= 0.0000
R-squared	= 0.2333
Adj R-squared	= 0.2325
Root MSE	= 18.439

radltw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	-1.098569	.7076211	-1.55	0.121	-2.487133	.2899956
_cons	54.56074	2.231642	24.45	0.000	50.18159	58.93989
rho	.8440384					

Continued on next page ....

Table 6 Female perception of health risk from cumulative exposure to radiation

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	123371.676	1	123371.676
Residual	405130.971	1087	372.705585
Total	528502.647	1088	485.756109

Number of obs	= 1089
F(1, 1087)	= 331.02
Prob > F	= 0.0000
R-squared	= 0.2334
Adj R-squared	= 0.2327
Root MSE	= 19.306

radltw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	1.590909	.7164974	2.22	0.027	.1850345	2.996784
_cons	60.47501	2.155678	28.05	0.000	56.24525	64.70477
rho	.8185211					

The Prais-Winston regression is needed because of the large autocorrelation in the dependent series, which is designated as ρ at the bottom of each table. In both the male and female table, the non-zero mean is statistically significant. The slope is not significant in the male table, but it is significant in the female table. This suggests that a first differencing of the female but not the male cumulative danger measure might be in order. The slope is not steep however, and a natural log transformation might render the dependent variable more tractable if the variance was proportional to the mean and this were the main source of the problem. However, a natural log alone, would not resolve this issue. Nevertheless, the slope is not very steep and we will try to control for this temporarily by including the time trend among the explanatory variables in the model to partial out its main deterministic effects.

2.3 The proportion of pollution attributable to Chornobyl

Next we examine the variable we designate as *radchw*, which is the proportion of pollution that the respondent attributes to Chornobyl radioactive fallout. We have observed that the males and females may have different perspectives on the nature of this pollution. Therefore, we graph the lowess plots of this measure over the three waves of our study and display those graphs in Figure 4.

Figure four shows that males report a declining percent of pollution due to Chornobyl, whereas the women report a more or less stable level of pollution due to Chornobyl. In wave one, the male level appears to be about 56 percent and by wave three, the percent appears to be around 53 or 54 percent. Is a decline of two percent over 30 years statistically significant? The female level begins at about 61 percent and ends at about the same level some thirty years later. The questions arises as to whether the male reports indicate a significant downward trend over the waves and as to whether the female reports indicate a stable level over the waves of this percent of pollution due to Chornobyl.

If we elaborate on the means from wave to wave, for the men, the percent of the health cumulatively affected goes from 53.09% in wave one, to 48.7% in

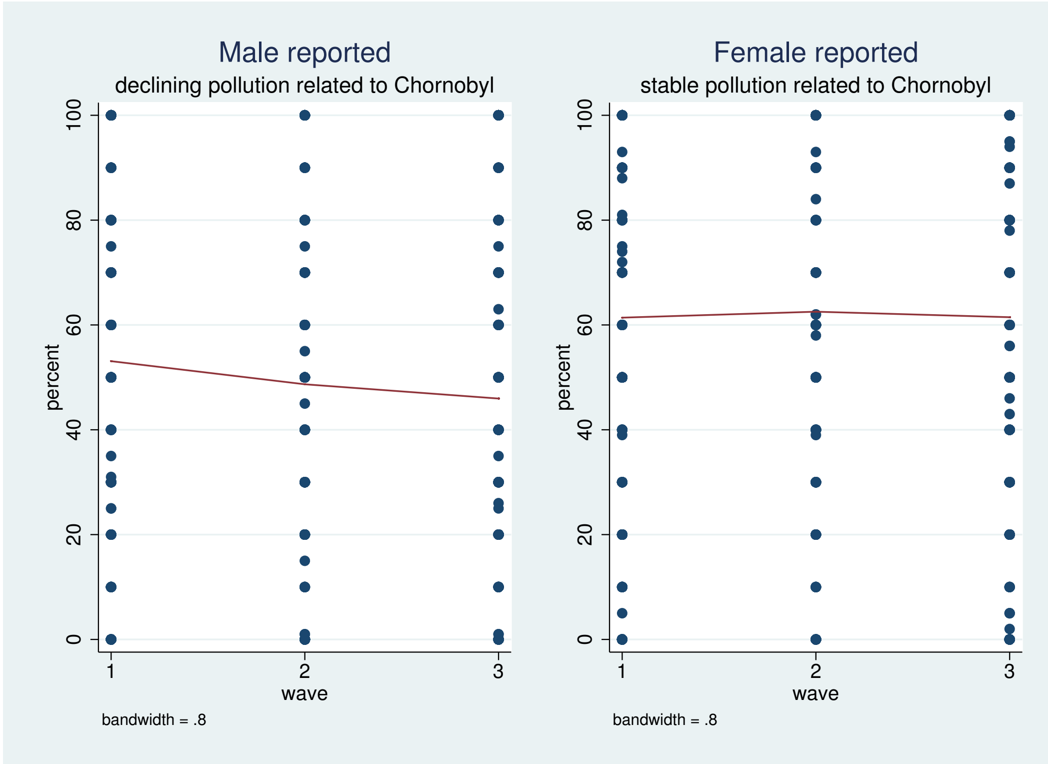

Figure 7: Male and Female reports of the percent of pollution attributable to Chornobyl

wave two, to 45.96% in wave three, which is evidence of a downward trend. From Table 7, we can observe that this trend based on the period or waves is statistically significant. In sum, the men believe that the clean-up is making progress and that pollution due to Chornobyl is declining.

The women do not seem to be so sure. In wave one, they report that 61.397% report of their health has been affected by cumulative exposure to radiation. A decade after Chornobyl, this level was raised to 62.53% by them. By the time of the interview, this estimated level dropped to 61.48%. However, the significance tests shown in Table 8 do not reveal that the slope of this broken trend exhibits any long-range significance. In brief, Table 7 shows that both the intercept and the slope of the males are statistically significant, whereas Table 8 reveals that only the constant in the female model is statistically significantly different from zero.

Whether for males or females, the trend appears to be more negative for the percent of pollution due to Chornobyl in the interim between waves two and three than in that between waves one and two. This is more pronounced in Figure 8 due to the different aspect ratio, but we have to resort to the statistics for an assessment of whether or not there is a statistically significant trend in that direction.

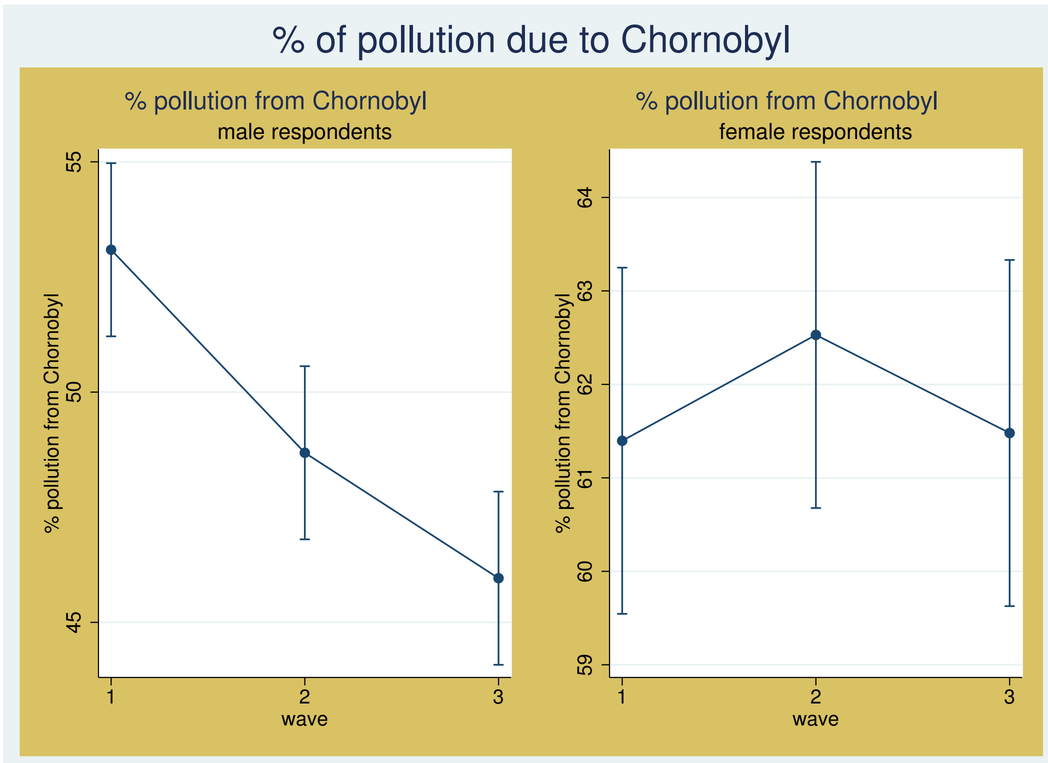

Figure 8: Male and Female reports of % pollution related to Chornobyl

Table 7 Male perception of percent of pollution due to Chornobyl

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	149540.24	1	149540.24
Residual	459088.055	1018	450.970584
Total	608628.295	1019	597.279975

Number of obs	= 1020
F(1, 1018)	= 331.60
Prob > F	= 0.0000
R-squared	= 0.2457
Adj R-squared	= 0.2450
Root MSE	= 21.236

radchw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	-3.566176	.814366	-4.38	0.000	-5.164205	-1.968148
_cons	56.57108	2.29398	24.66	0.000	52.06961	61.07255
rho	.7718814					

We observe that the negative trend for males is statistically significant, but the positive trend for females is not. Whether the analysis is performed for males or females, the autocorrelation coefficient is smaller than before and less likely to be confused with a unit root than in the previous analyses. Our use of wave among the explanatory variables tends to control for any fixed aspect of that trend.

Table 8 Female reports of percent of pollution due to Chornobyl

.

Prais-Winsten AR(1) regression -- iterated estimates

Source	SS	df	MS
Model	148550.509	1	148550.509
Residual	493927.245	1087	454.39489
Total	642477.754	1088	590.512642

Number of obs	= 1089
F(1, 1087)	= 326.92
Prob > F	= 0.0000
R-squared	= 0.2312
Adj R-squared	= 0.2305
Root MSE	= 21.317

radchw	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wave	.0413223	.791131	0.05	0.958	-1.510994	1.593639
_cons	61.46439	2.245088	27.38	0.000	57.05919	65.86959
rho	.7779387					

2.4 Recapitulation

In answer to the question of whether there is evidence of a direct dose-response relationship as evidenced by trends in risk perception we can say that there is preliminary evidence of a significant positive trend in the self-reported Chornobyl-related health-risk perception on the part of females in Kiev and Zhitomyr Oblasts of the Ukraine, as shown in Tables one and two, as well as in Figures one and two above. Although the initial level of the percent of health affected is statistically significant for both men and women, only the slope of the females exhibits statistical significance. Nevertheless, both trends are positive and that of the women is steeper than that of the men. The size of the confidence intervals makes it difficult if not impossible to maintain that the male trend is statistically significant, whereas that of the female is according to the Prais-Winston regression analysis output in Table two. Neither autocorrelation parameter. ϕ_1, is large enough to be confused with a unit root. For the males, $\phi_1 = 0.878$ and for the females, it is equal to 0.806.

When we examine the relationship between health hazard to the family posed by the Chornobyl radiation, the results are similar. Figures three and four reveal broken trends where the initial levels are both statistically significant from zero. The slopes become less steep over time for both males and females. Nonetheless, the starting values for the females are higher than those for the males and the overall slope is steeper for the females as well. However, in this case, both trends are statistically significant and positive. Neither autocorrelation parameter, ϕ_1 is high enough to be confused with a unit root. For the males it was 0.818, whereas for the females, $\phi_1 = 0.0.69$

3 Dose-Psychological Response relationships

3.1 Risk perceptions as a function of dose

When we consider risk perceptions as a function of dose, we do so in a panel analysis with panel corrected standard errors. The corrections are performed for heteroskedastic panels and autocorrelation. In these models, we not only control for wave, but also for distance from Chornobyl, time-varying age, and for a categorical effect for raion.

When we observe the broken trend and leveling of the self-perceived Chornobyl risk and observe that the change is not statistically significant, particularly from wave two to wave three, we may appreciate that it may be difficult to show a linear relationship between other variables and that sense of risk. However, we do not know whether that will be the case until we attempt to do this with the proper controls. When we run a repeated measures MANOVA on the age of the respondent, the distance from Chornobyl in kilometers, the square of that distance, the raion of residence, and the average cumulative dose in wave three, we do find evidence of statistically significant relationship between dose and Chornobyl related family health risk. For the males, this effect has Roy's largest root at 0.0411, with F(df1=3.0, df2=255.0)=3.49, with a significance level= 0.0163. For the females, Roy's largest root = 0.0295 F(df1=3.0,df2=256.0)=2.52, with a significance level= 0.0586. It may be that we need to apply more controls. Surely, the preliminary results are tantalizingly close to suggesting that with a more elaborate analysis, we may learn much more than what we have found thus far.

If we subdivide our sample into two groups, based on the proximity of their residences to Chornobyl, do we find *prima facie* evidence for suspecting that there might be a Risk-distance relationship. We construct a variable called `proximity`, indicating whether or not a respondent lives 20 miles or less from the Chornobyl accident site. We find that five males report living within that range and six females. Then we preform a Mann-Whitney sum of ranks test, which does not depend on the asymptotic nature of the distribution, to determine whether there is a significant difference in Self-perceived or self-reported familial Chornobyl related risk, and we find that there appears to be a significant difference between the male reports of those who live less than 20 miles from Chornobyl and those who live farther away in the percent of risk to the family. Although the self-reported health risk is almost but not quite statistically significant, none of the female responses exhibited a statistically significant difference according to this test.

```
Table 8a Male reports of self-perceived Chornobyl health risk

. ranksum radhlw3 if gender==1, by(proximity)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
   proximity |      obs    rank sum    expected
-------------+---------------------------------
21 + miles a |      335     56752.5     57117.5
less than 20 |        5      1217.5       852.5
-------------+---------------------------------
    combined |      340       57970       57970
unadjusted variance    47597.92
adjustment for ties     -768.97
                     ----------
adjusted variance      46828.95
Ho: radhlw3(proxim~y==21 + miles away) = radhlw3(proxim~y==less than 20 miles)
             z =  -1.687
    Prob > |z| =   0.0917
```

```
Table 8b Male report of Chornobyl related risk to family

. ranksum radfmw3 if gender==1, by(proximity)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
   proximity |      obs    rank sum    expected
-------------+---------------------------------
21 + miles a |      335       56669     57117.5
less than 20 |        5        1301       852.5
-------------+---------------------------------
    combined |      340       57970       57970
unadjusted variance    47597.92
adjustment for ties    -1392.77
                     ----------
adjusted variance      46205.15
Ho: radfmw3(proxim~y==21 + miles away) = radfmw3(proxim~y==less than 20 miles)
             z =  -2.086
    Prob > |z| =   0.0369
```

Although none of the other reported risk perceptions exhibited significant differences according to the residential proximity to the Chornobyl site, these indications suggest further exploration of the trends in risk perception are warranted.

3.2 Reconstructed dose and Nottingham health profile subscales

We had hypothesized that there would be a dose response relationship with respect to psychological symptoms. To discover whether this is so, we will begin to examine the relationship between the average cumulative reconstructed dose of a key component of the fallout and some well known and widely used psychological scales and sub-scales. We measure $Cesium^{137,}$ a radioactive isotope that has a half-life of approximately 30 years, in μGrays. This principle part of the radioactive fallout that was given off by Chornobyl will be used as an independent variable, embedded with other potential confounding variables such as age, wave, gender, along with computed geodesic distance from Chornobyl (based on spherical trigonometric computations, using the Thaddeus Vincenty formula for the distance between the residence of the respondent and the accident site, as well as a categorical variable for the raion difference.

When we control for age, distance from Chornobyl, distance from Chornobyl squared, the raion in which the respondent lives, and test the relationship of all of the Part 1 of the Nottingham Health profile against the cumulative dose, even at wave three which is when it would be largest, we do not find a significant effect for males or for women. For men, the effect as measured by Roy's largest root (Roy's largest root=0.0452, F(1,252) = 1.90, p=0.082). For women, Roy's Largest Root = 0.222, F(6, 253) = 0.94, p=0.47).

When we examine each of the items Weighted Part I items of the Nottingham Health Profile for preliminary evidence an effect of cumulative dose on those subscale, we control for distance from Chornobyl and its square, as well as the raion of residence. For the males, we observe evidence of an effect on energy level (p=0.28) and almost some evidence of an effect on emotional reaction (p=0.054). The female respondents, we note that there is almost evidence of an effect on social isolation (p= 0.076), but otherwise we find no preliminary evidence of a relationship.

Part 2 of the Nottingham health profile consists of a battery of binary coded items, for which reason we employ a logistic regression analysis on such items to test whether there might be some impact. We control for age, distance from Chornobyl, distance from Chornobyl squared, the raion of current residence, and examine the logistic regression coefficients in Table 9 below for the logistic regression coefficients for cumulative dose at wave three and their association with the binary coded items in Part 2 of Nottingham Health Profile, listed in the leftmost column of the table.

Table 9 Logistic Regressions of cumulative dose at wave 3 on Part 2 of Nottingham Health Profile

Dep. var.	b	se	z	p
Male				
Paid emplymt	.0099	.0441672	0.22	0.823
Homecare	-.0009	.0463303	-0.02	0.984
Social probs	.0333	.0442964	0.75	0.452
Fam probs	-.0897	.1795068	-0.50	0.617
sex life	.0449	.0427661	1.05	0.294
ints-hobbies	-.0591	.0989439	-0.60	0.550
vacation	.0017	.0565368	0.03	0.976
Female				
Paid emplymt	0.088	0.0649	1.58	0.128
Homecare	-0.097	0.7752	-1.25	0.211
Social probs	0.321	0.0953	3.36	0.001**
Fam probs	0.063	0.0795	0.79	0.427
sex life	0.170	0.0765	2.23	0.026*
ints-hobbies	0.111	0.0695	1.60	0.110
vacation	0.129	0.0712	1.80	0.069

Legend: * = p < .05, ** = p < 0.01, and *** = p < 0.001

To ascertain whether there was any dose response on these items we performed a logistic regression analysis. Although we obtain no evidence of an impact on the lives of the men, we note that with respect to social problems and sex life on the part of the women, there is evidence of an effect. Given the potential of radioactivity for impacting the reproductive organs, this finding is not surprising.

As an endogenous variable in the model, we employ the separate Nottingham health profile sub-scales. Part 1 of the Nottingham health profile consists of weighted items such as the extent to which Chornobyl affected the energy level, the emotional reaction, the ability to sleep, the social isolation, the pain, or the physical ability of the respondent. Part 2 of the scale consists of dummy variable responses as to what aspects of a respondent's life was affected– paid employment, home care, social life, family problems at home, sex life, interests and hobbies, or vacation plans. Because the items on part 1 and part 2 are highly correlated. On part 1, the correlation between energy level and emotional reaction is .5064 for males and 0.60 for females. The correlation between energy level and physical ability for males is 0.51 and .54 for females. On part 2, there are several correlations in the .50-.60 range for males and for females. Therefore, one approach to analyzing them would be to see whether dose had a significant effect on them, when they are taken altogether.

The statistical technique employed to test the hypothesis is a Prais-Winston autoregression model with panel corrected standard errors, such that we allow for specific autocorrelation peculiar to each panel. From such a technique we have indication that there may be a statistically significant direct relationship between average cumulative dose of CS^{137} at each wave with energy level for men and women, emotional reaction for men and women, sleep for both men and women, social isolation for women, pain for both men and women, as well as a direct relationship with physical ability for women. The regression coefficients for the average cumulative dose with respect to the dependent variables of the Nottingham Health Profile are shown in Table 10.

Table 10 Preliminary Indications of dose-psychological response on Nottingham Health Profile

Regression Coefficients for avgcumdosew

Part I

Dep. var.	b	se	z	p
Male				
Energy Level	1.563456	.5732973	2.73	0.006**
Emtnl Reactn	.4946091	.0807462	6.13	0.000***
Sleep	.7797975	.2823896	2.76	0.006**
Social Isola	.2284929	.1377598	1.66	0.097
Pain	.4560564	.2178293	2.09	0.036*
Physical abil	-.0654486	.1639261	-0.40	0.690
Female				
Energy Level	1.389362	.396265	3.51	0.000***
Emtnl Reactn	.7032247	.26407	2.66	0.008**
Sleep	1.658802	.4410328	3.76	0.000***
Social Isola	.910336	.29499	3.09	0.002**
Pain	1.141051	.4426512	2.58	0.010**
Physical abil	.9487973	.2463051	3.85	0.000***

Legend: * = p < .05, ** = p < 0.01, and *** = p < 0.001

3.3 Dose reconstruction and the Basic Symptom Inventory

Table 11 reveals the preliminary indications of dose-psychological response relationship when the impact is measured by the Basic Symptom Inventory subscales. From this table, there appears to be significant association between dosage on the one hand and psychological symptomatology exhibited by the individual on the other. The spectrum of such assocation appears to be broader with respect to the women than it does with respect to the men.

Table 11 Preliminary dose-psychological response

indicated by panel regression coefficients of avgcumdosew on BSI

Dependent variable
BSI scale/subscale

Male model	b	se	z	p
BSI total	.397032	.2058957	1.93	0.054
Positive symp	.388887	.192418	2.02	0.043*
Global svrty	.0074912	.0038848	1.93	0.054
BSI somatic	.0290447	.0428616	0.68	0.498
Obsesive-cmp	.0821425	.0253838	3.24	0.001**
Intrpsnl sens	.0211769	.025334	0.84	0.403
Depression	.021425	.0257616	0.83	0.406
Anxiety	.0491167	.0221589	2.22	0.027*
Phobic anx	-.0447182	.0115933	-3.86	0.000***
Hostility	.0728123	.0313953	2.32	0.020*
Paranoia	.0633425	.0142001	4.46	0.000***
Psychoticism	.0258878	.0252032	1.03	0.304

Female model	b	se	z	p
BSI total	2.590682	.4899928	5.29	0.000***
Positive symp	2.473533	.4797989	5.16	0.000***
Global sevrty	.0488808	.0092451	5.29	0.000***
BSI somatic	.2680485	.0710858	3.77	0.000***
Obsessive-cmp	.3503427	.0676711	5.18	0.000***
Intrpsnl sens	.3222848	.0610448	5.28	0.000***
Depression	.2852299	.0750644	3.80	0.000***
Anxiety	.3385903	.0742483	4.56	0.000***
Phobic anx	.2586442	.061283	4.22	0.000***
Hostility	.0051921	.030978	0.17	0.867
Paranoia	.1621083	.0761346	2.13	0.033*
Psychoticism	.2840319	.0471464	6.02	0.000***

Legend: * = p < .05, ** = p < 0.01, and *** = p < 0.001

3.4 Reconstructed dose, coping, and Post-traumatic stress disorder

We also examine the relationship between dose and coping response as well as any evidence for PTSD following the nuclear disaster. By controlling for time, distance from the Chornobyl accident site, the particular raion of residence, as well as the wave in which the analysis is performed, a Prais-Winston regression, using standard errors corrected for heteroskedasticity and first-order panel specific autocorrelation, reveals a direct relationship between reconstructed dose, measured in μGrays, of $Cesium^{137}$ and coping (by avoidance) as well as PTSD for women. However, no such relationship is observed when the same analysis is performed on male respondents. When the dependent variable is the coping subscale or the Mississippi PTSD scale in the first column of Table 11 and the average cumulative dose of CS^{137} is the independent variable, the panel parameter estimates for average cumulative dose are listed in Table 12.

Table 12 Preliminary dose-psychological response

indicated by panel regression coefficients of avgcumdosew on PTSD and Coping Scales

Dependent variable
PTSD scale or Coping subscale

Male model	b	se	z	p
MiPTSD	.0807494	.058857	1.37	0.170
CSprbslv	.0146769	.0360496	0.41	0.684
CSsocspt	.0711428	.0502657	1.42	0.157
CSavoid	-.0317371	.0187768	-1.69	0.091
Female model	**b**	**se**	**z**	**p**
MiPTSD	.6819366	.209325	3.26	0.001**
CSprbslv	0.048971	.0436565	1.12	0.262
CSsocspt	-.0513008	.0495974	-1.03	0.301
CSavoid	.0910715	.045783	1.99	0.047*

4 Reconstructed dose and the medical diagnosis

The question naturally arises about whether the illnesses diagnosed by the physicians stem from the risk perception or whether they derive directly from the exposure on the part of the respondents to the radiation.

To answer this question, we controlled for the same variables mentioned in the coping and PTSD panel analysis with the same type of panel data model. We found that a significant relationship appeared between self-perceived Chornobyl risk and several medically diagnosed illnesses. The male model output is listed in Table 14 and the model output for females is displayed in Table 15 that follows.

The medically diagnosed illnesses indicated as ICD9 codes reported by male respondents that emerged as statistically significant included the following diseases. Not all parameter estimates were positive. Only the code for hypertension exhibited a direct statistically significant positive relationship with self-perceived Chornobyl-related health risk. All of the others were significantly negatively related to such health risk for the males, with the exception of gastritis, which faded from significance as the model was pruned of its nonsignificant effects. If the ICD9 coded illnesses were specifically related to other sources of tension and anxiety, their parameter estimates might be negative while the parameter estimates of average cumulative reconstructed dose and self-reported Chornobyl related health risk stands out as being significantly positive.

Table 13 International Classification of Disease (ICD9) code

variable name	variable label
icdx1nr5	icdx1nr==401 hypertension
icdx1nr15	icdx1nr==542 other appendicitis
icdx3nr8	icdx3nr==ac bronchitis/brnchial
icdx4nr20	icdx4nr==osteochondropathies
icdx5nr11	icdx5nr==gastritis/duodenitis
icdx6nr17	icdx6nr==oth inflamm polyarthrop

Table 14 A Panel regression with panel corrected standard errors for male respondents

Prais-Winsten regression, heteroskedastic panels corrected standard errors

Group variable:	id	Number of obs =	945
Time variable:	wave	Number of groups =	315
Panels:	heteroskedastic (balanced)	Obs per group: min =	3
Autocorrelation:	panel-specific AR(1)	avg =	3
		max =	3
Estimated covariances =	315	R-squared =	0.6968
Estimated autocorrelations =	315	Wald chi2(10) =	453.36
Estimated coefficients =	11	Prob > chi2 =	0.0000

radhlw	Coef.	Het-corrected Std. Err.	z	P>|z|	[95% Conf.	Interval]
magew	.0911078	.0777407	1.17	0.241	-.0612613	.2434768
wave	-.3105179	.8473812	-0.37	0.714	-1.971354	1.350319
HavKm	.000331	.0002739	1.21	0.227	-.0002058	.0008678
avgcumdosew	1.046099	.3281148	3.19	0.001	.4030058	1.689192
icdx1nr5	24.27428	8.120844	2.99	0.003	8.357714	40.19084
icdx1nr15	-35.01613	2.721414	-12.87	0.000	-40.35	-29.68225
icdx3nr8	-39.17173	5.099995	-7.68	0.000	-49.16753	-29.17592
icdx4nr20	-26.493	11.0432	-2.40	0.016	-48.13727	-4.848726
icdx5nr11	-6.887478	4.798842	-1.44	0.151	-16.29304	2.51808
icdx6nr17	-31.90259	3.911709	-8.16	0.000	-39.5694	-24.23578
_cons	42.65546	4.928691	8.65	0.000	32.9954	52.31551
rhos =	.7534247	1	1	.9883219	.9725246 ...	.9852289

In Table 14, radhlw is the self-reported Chornobyl-related health risk in percent, whereas magew is the time-varying age of the respondent at the wave under consideration. ρ, the autocorrelation coefficient is not so large as to be confused with a unit root, which could result in infinite and non-estimable variances, as well as a breakdown of the computation process. Wave is the particular wave under consideration, which serves as a control for the fixed or deterministic wave trend effect. HavKm is the number of kilometers from Chornobyl as computed by the Haversine formula for geodesic distances. Avgcumdosew is the average cumulative reconstructed dose of CS^{137} in microGrays, while *_cons* is the constant in the equation.

Table 15 Interational Classification of Disease (ICD9) code
variable name variable label

variable name	variable label
icdx1nr6	icdx1nr==410 myocardial infarction
icdx4nr5	icdx4nr==rheum fev w/o hrt involv
icdx4nr13	icdx4nr==regional enteritis
icdx4nr20	icdx4nr==osteochondropathies
icdx5nr11	icdx5nr==gastritis/duodenitis
icdx6nr17	icdx6nr==oth inflamm polyarthrop

The model for the women is a little more inclusive than that for the men. It contains in addition to the aforementioned ICD9 codes a few others. Listed in Table 15, they include heart attack, other dermatitis, rheumatic fever, regional enteritis, and bone diseases, while stroke faded from statistical significance as the model was pruned.

Table 16 Panel regression with panel corrected standard errors for female respondents

```
Prais-Winsten regression, heteroskedastic panels corrected standard errors
Group variable:   id                            Number of obs      =      1035
Time variable:    wave                          Number of groups   =       345
Panels:           heteroskedastic (balanced)    Obs per group: min =         3
Autocorrelation:  panel-specific AR(1)                         avg =         3
                                                               max =         3
Estimated covariances      =       345          R-squared          =    0.6433
Estimated autocorrelations =       345          Wald chi2(13)      =   1208.94
Estimated coefficients     =        14          Prob > chi2        =    0.0000
```

radhlw	Coef.	Het-corrected Std. Err.	z	P>|z|	[95% Conf.	Interval]
magew	.4486733	.1384132	3.24	0.001	.1773885	.7199581
wave	-2.392871	1.352742	-1.77	0.077	-5.044197	.2584553
HavKm	-.0003265	.0002564	-1.27	0.203	-.000829	.000176
avgcumdosew	2.326189	.740102	3.14	0.002	.8756154	3.776762
icdx1nr5	7.824671	5.91305	1.32	0.186	-3.764694	19.41404
icdx1nr6	-25.15839	10.71269	-2.35	0.019	-46.15489	-4.1619
icdx1nr15	-9.560759	10.73448	-0.89	0.373	-30.59995	11.47843
icdx1nr17	-22.08058	8.348168	-2.64	0.008	-38.44269	-5.718471
icdx4nr5	41.79322	3.675903	11.37	0.000	34.58858	48.99786
icdx4nr13	-21.1469	2.783767	-7.60	0.000	-26.60299	-15.69082
icdx4nr20	-25.79742	10.0804	-2.56	0.010	-45.55464	-6.040201
icdx5nr11	-25.93485	4.245408	-6.11	0.000	-34.2557	-17.614
icdx6nr17	-42.28564	3.201526	-13.21	0.000	-48.56051	-36.01076
_cons	50.54379	5.014499	10.08	0.000	40.71555	60.37203
rhos =	-.3568529	.9545388	1	1	.7441786 ...	1

These results are preliminary ones. Although attempts to regenerate significant results in the reverse direction with a panel logistic regression model using the medical diagnoses as dependent variables came to yielded no statistically significant relationships, which indicates that the directionality of these relationships is from the medical diagnoses to the self-reported Chornobyl related health risk and, with the use of the same control variables, not the other way around.

However, it is possible that the application of different control variables for other types of analysis might minimize multicollinearity, generating somewhat different results. Whereas we only used wave as a preliminary control variable to control for the deterministic trend in the first set of equations, we might apply additional control variables that we used in the last few equations. The results could change somewhat if we applied those controls. For example, if we set out to use the model to forecast, we might split the sample, estimate on one part, and validate on the other. Under those circumstances, we might be less tolerant of residuals correlated with the regressors and we could we could applying tighter controls obtain slightly different results. If we perform policy analysis, we ought to use higher standards for exogeneity, requiring superexogeneity, than we do for modeling or forecasting [5]. Under such circumstances as those, we might use additional control variables, stressing parameter constancy, which could lead to different models altogether. Moreover, we may try to robustify our models further to provide for less fragility and more reproducibility by taking steps to trim or downweight outliers. We might try to bootstrap these results to establish them further[7]. Were we to try to optimize the fit of a regression model, we might apply the techniques of AutoMetrics, although whether this approach would be optimal for specific hypothesis testing in this case is open to debate several other approaches that we may take before we settle upon the models as being the ones we consider optimal[4],[2],[1], and [8]. For these reasons, the results we find our deemed preliminary and subject to further testing. In the meantime, these results provide preliminary evidence for the hypotheses that we have considered, and thereby offer some sense as to which variables may be instrumental in the formation etiological pathways of psychological symptomatology following a nuclear disaster.

References

[1] Castle, J., Doornik, J., Hendry, D.F. 2011 Evaluating automatic model selection *Journal of Time Series Econometrics*,**1**,3, 1-30. Berkeley Electronic Press, 2011.

[2] Jurgen Doornik and Hendry, Sir.D.F. 2009 *PcGive 13: Vol 1. Empirical Econometric Modeling* London, UK: Timberlake Consultants, Ltd., 166-180.

[3] Elashoff, J, 2007 *NQuery Advisor*, version 7.0. Cork, Ireland:Statistical Solutions, Ltd.

[4] Engle, R., Hendry, D.F., and Richard, J-F. 1994 *Exogeneity* Ericsson, N.R. and Irons, J. eds. *Testing Exogeneity* Oxford, UK: Oxford University Press, 36-72.

[5] Ericsson, Neil R. *Introduction Testing Exogeneity* Oxford, UK: Oxford University Press, 7-19.

[6] Granger, C.W.J. and Newbold, P. *Forecasting in Business and Economics*, 2nd. ed. San Diego, CA: Academic Press, 130.

[7] Harrell, Jr., Frank E. 2001 em Regression Modeling Strategies New York: Springer, 1-80, 426-427.

[8] Pindyck, R. and Rubinfeld, D. 1997 *Economic Models and Economic Forecasts* New York: McGraw Hill, 345-365.

[9] Pitt, M.K. and Shephard, N. 1997 *Filtering via Simulation: Auxiliary Particle Filters* Journal of the American Statistical Assocation Vol. 94, June 1999, 590-599.