EXHIBIT C
FILED UNDER SEAL

Beate Ritz

Paraquat Report June 2020

1. Beate Ritz, MD, PhD, Background and Qualifications

I, Beate Ritz, MD, Ph.D., am Professor of Epidemiology at the UCLA Fielding School of Public Health, and served formerly as Vice Chair and Chair of the Epidemiology Department for a decade, and hold co-appointments in Environmental Health Sciences and Neurology at the UCLA, School of Medicine. I was trained in Medicine at the University of Hamburg/Germany and received a doctoral degree from the University of Hamburg in Medical Sociology in 1986. I furthermore received another doctoral degree in Epidemiology from UCLA in 1995, and subsequently was hired as a faculty member at UCLA. My faculty appointment at UCLA is one of several positions specifically assigned to the Center of Occupational and Environmental Health (COEH) mandated by the State of California to conduct research, teaching, and service to communities in California on occupational and environmental health issues. Hence, my primary research interests are health effects from occupational and environmental exposures with a focus on pesticides and air pollution and chronic diseases including cancers, reproductive outcomes, neurodevelopmental disorders and neurodegenerative diseases. I served for more than a decade as the co-director of the NIEHS-funded UCLA Center for Gene-Environment Studies in Parkinson's disease (PD) and am currently the Director of the American Parkinson's Disease Association Center for Excellence in PD Research. In the past two decades, I was the principal investigator of numerous Parkinson's disease, pesticides and gene-environment epidemiology studies in California and also conducted research based on large databases (such as cancer registries) assembled in California and Denmark. As part of my research, I developed geographic information system (GIS) based exposure assessment tools to assess chronic health effects of long-term pesticide exposures and of air pollution in California. In the early 2000s, I served as a member of the external advisory committee for the NCI/NIEHS Agricultural Health Cohort Study and for one year chaired this committee. I also was a visiting scientist at IARC/Lyon in 2006-07. In 2007, I received the Robert M. Zweig M.D. Memorial Award (Clean Air Award) from the California South Coast Air Quality Management District and in 2008 I was awarded the "Excellence in Research" award from the American Parkinson's Disease Association. I served on multiple National Academy of Sciences/Institute of Medicine (NAS/IOM) committees evaluating Gulf War Illness – including IOM reviews of cancer and of amyotrophic lateral sclerosis (ALS). Recently, I served on the NAS/IOM committee on "Incorporating 21st Century

Science into Risk-Based Evaluations" and I just newly began serving on the committee to assess "Health Effects in Vietnam Veterans from Agent Orange (herbicides)". I am a CA Governor appointed member of the scientific review board for the California Air Resources Board (CARB) panel on Air Toxics. I served on the editorial Board of the Journal *Epidemiology* as well as other journals (currently I am editing a section of the journal *Current Environmental Health Reports*). I have been the elected President Elect of the International Society for Environmental Epidemiology (ISEE) in 2018-19, and currently serve the Society as Past President. In 2018, I have been named as being in the top 1% of most cited authors in my field by Clarivate Analytics. My Curriculum Vitae is attached.

2. Methodology

2.0 Definitions of terms.

A Cohort study is usually a special group of people such as employees of one or several large companies (automobile manufacturers, mines etc) or members of a profession (such as nurses, teachers, or farmers) or a religious group (e.g. Seventh Adventists). A cohort study enrolls subjects unaffected by the disease of interest (here Parkinson's disease) and assesses their exposure status (in this case, to paraquat) at baseline, and then follows this special group of people over time to determine who develops the disease(s) of interest and - if possible - to update exposure information over time. At enrollment and - if possible during follow-up - all participants are asked to report their past exposures, or exposure is reconstructed from records or job titles (see also below under case-control study). Note that at time of enrollment no study participant is allowed to suffer from the disease of interest yet, i.e., at the time of baseline exposure assessment. In many cohorts, exposure is only assessed at enrollment (baseline) which is problematic if follow-up is long and exposures continually change until disease occurs. For rare diseases (with less than 10% prevalence in the population), such as Parkinson's disease, we generally would have to follow a very large group of people (ideally 50-100,000) over a very long time (10-20 years) to accrue a sufficient number of cases and make statistical analyses informative enough to reach conclusions.

A (Population-based) Case-control study assembles its subjects by disease status. That is, study subjects referred to as cases are enrolled because they already have the disease (in this case, PD)

and controls are those who at the time the cases were diagnosed do not have the disease of interest. A study is considered 'population-based' if the controls are being selected from the same population that gave rise to the cases enrolled in the study. Everyone enrolled is either asked to report their past exposures (in this case, paraquat) in interviews or on questionnaires or – if possible – exposures may be reconstructed from a record system (e.g. sales records or pesticide application records) or by experts who evaluate job tasks and titles reported by all study participants (this method is generally referred to as creating a job exposure matrix).

Odds Ratio (OR). An odds ratio, or OR, is considered one measure of association between an exposure and a disease similar to risk or rate ratios (see below). The OR represents the odds that the disease will occur in a group of people with a particular exposure, in comparison to the odds of the disease in a group of people without this exposure. An OR of 1.00 is what we call the null value, meaning no effect as the odds of disease in the exposed equals the odds in the unexposed, and hence the ratio of the two values is 1. Thus, an OR of 1.40 as reported in one of the studies below, for example, represents a 40% increase in the odds of PD from exposure to paraquat. An OR of 3.10 as reported in one of the studies below represents a 210% increase in the odds of PD from the exposure. An odds ratio is a "point estimate" or the "central" estimate of the relationship between exposure and disease, in a given study (note: the OR is in the center of the upper and lower confidence limit boundaries, see below). Odds Ratios are the statistics that are used most often to analyze case-control studies, and they are often calculated using a statistical analysis tool called logistic regression; however, they also can be derived by simple calculations as a product ratio based on a 2x2 table of data on exposed and unexposed cases and controls. ORs like RRs (see below) per se do not represent a measure of clinical or public health impact; that is a 15% or 1.15 fold increase (OR of 1.15) in such an estimate has a different implication when the outcome is a minor skin mole versus a severe birth defect such that an exposure that adds 15 additional neural tube defects to every 100 such outcomes (15% excess risk due to exposure) will likely be considered actionable by parents and policy makers.

Rate Ratio (RR). A rate ratio is the measure of association between exposure and disease that can be calculated with cohort study data. The incidence rate we calculate here takes not only the number but also the time each study subject spends under observation into account; i.e. person-

time (number of years a subject spent under observation) depends on how much time has passed from the start of the study until the point in time when disease is diagnosed or until the end of follow-up of a study participant who never develops the disease such as the end of the study or him/her dropping out of the study. Thus, we compare the incidence rates of disease in those exposed (number of cases divided by the person-time of all subjects in the exposed group) to the incidence rate of disease among study participants who are not exposed. Therefore, a RR different from an OR inherently relies on measures that included time under observation (i.e. rates). However, the results are interpreted in the same way: a RR of 1.00 is the null (no effect); a RR of 1.50 is a 50% increase in the rate of disease among the exposed, etc.

Risk Ratio (or Relative Risk) is a ratio of the risk in the exposed subjects divided by the risk in the unexposed members of a cohort - where risks are defined as the number of (un)exposed cases divided by the total number of (un)exposed. Thus, different from rate ratios, this risk measure uses the number of subjects rather than the number of person-years subjects contribute during follow-up as the denominator. This method is used for well-defined (similar length) follow-up periods in the exposed and the unexposed such that the time under observation will not contribute additional information and we can substitute persons for person-time.

NOTE: under certain circumstances that are often met, especially for rare diseases, the odds ratio (OR), risk ratio (RR) and rate ratio (RR) are the same (albeit calculated as the ratio of odds, risks, or rates) and the interpretation of the estimates is also the same.

P-value. The p-value is the probability of obtaining an estimate at least as far from a pre-specified value (in case of the null hypothesis the 'null' value) as the estimate we have obtained, if the specified value were the true value. For example, a p-value of 0.05 means that, given the null hypothesis is true, if one repeatedly conducted 100 tests of samples drawn from the same population (people), then in 5% of your tests, one would obtain the results one got solely due to random error (chance). Note: no p-value, for the null hypothesis or any other hypothesis, is the probability that the specified hypothesis is true. It is a metric intended to show the likelihood of random error contributing to the results. It *should not* be interpreted as the probability that an agent causes an outcome.

Confidence interval (CI). A confidence interval, or CI, is the interval around an OR or a RR that likely includes the unobservable true measure of effect. In other words, it is an interval estimate (as compared to a point estimate) of the true underlying relationship between exposure and disease, in a given study. In practice, most estimates in scientific publications are 95% confidence intervals, which means that in 95 out of 100 times when sampling your study subjects, you will find the true result (effect estimate) within the given confidence interval.

N (number). The number of people in a study.

Statistical power is the ability of a study to estimate an effect size statistically with the data in hand. In essence, it is a reflection of the sample size (number of subjects in a study – in cohorts also the number of cases), the prevalence of exposure, and the expected effect size. Large sample sizes generally have higher statistical power, which means they have narrower and more stable confidence intervals around point estimates. Smaller sample sizes generally have wider confidence intervals, however, the statistical power also depends on the optimal exposure distribution and contrast and the expected size of the effect such as a RR or OR. Thus, larger studies are more powerful in finding statistically significant results when exposures or outcomes are rare especially when the expected size of the effect is moderate or small.

Data pooling or pooled analysis. To pool data is to use the raw (un-analyzed or non-summarized) data from several studies and merge them together to conduct analyses. Data pooling may be done when multiple studies have been done on a topic but were too small to conduct some informative subgroup analyses such as by gender or race etc. Data pooling increases the overall sample sizes and allows for a uniform approach to the analysis of the data but requires that the same or similar enough data have been collected in all studies such that it can be combined. Pooled studies have greater statistical power than the original studies from which they draw data but require considerable effort to "harmonize" (make similar) all the data. In order to conduct data pooling, scientists also need to have permission to access individual level data from the investigators of multiple studies.

Meta-analysis. Often, it might be scientifically of interest to pool data and summarize across studies, but scientists do not easily have access to the raw data from individual studies, or these studies were conducted in such different manners (e.g. in design or data collected) that the data cannot be pooled. This is because some studies may have been conducted decades earlier and the investigators cannot be reached or do not have access to the raw data anymore, or perhaps because investigators do not know and trust others, or they have to adhere to human subject restrictions (privacy laws for instance) that do not allow for the sharing of raw data. Thus, it is quicker and more efficient as well as possibly more appropriate to conduct a meta-analysis based on summary estimates from published and peer-reviewed reports. A meta-analysis uses the Odds Ratios or Rate Ratios and confidence intervals published in the original studies and combines them into a summary estimate for the estimated effect of the exposure on disease overall. Similar to pooled analyses, meta-analyses also have much greater statistical power than each study alone, but the authors do not have the option of re-analyzing the original data as might be possible if raw data were available from all original studies (such as lagging exposures or generating different exposure categories or adjusting for confounding variables etc.). However, there are many choices to be made for meta-analysis, the first and foremost of which is the studies to select or for which studies it is appropriate to generate a summary estimate. Even though it is possible to generate summary estimates across all studies published, this might not be appropriate or the most informative way to assess the literature. For example, a simple summary estimate across all studies may not make important distinctions between study types and might even be less scientifically informative than exploring what is called the "heterogeneity" of effect (risk) estimates across certain categories of studies such as case-control versus cohort studies, or studies using different exposure assessment methods (self-reported versus recorded exposures) or types of exposures (environmental vs occupational). By using the later approach, it is possible to learn from differences in study approaches and results; e.g., if we believe that occupational exposures are higher on average than environmental exposures, we would expect to see larger effect sizes in studies of occupational exposures. Thus, comparing the effect estimates of occupational and environmental exposure studies is one possible way to investigate whether there might be a dose response even if each individual study only assessed one exposure.

Null hypothesis means no effect is expected. In the studies described below, the null hypothesis generally was that PD is not related to paraquat exposure. The statistical tests done in the studies described below aim to test the null hypothesis: they aim to determine whether the null hypothesis can be rejected with adequate statistical certainty and whether any associations between exposure to paraquat and the development of PD is suggested by a study.

Dose-response. A dose-response association represents an increasing risk with an increasing exposure or dose, such that a larger number of days per year, or a longer number of years, or higher intensity of exposure is related to higher Odds Ratios. For example, the overall study Odds Ratio might be 1.40, but for people who used paraquat more often, the Odds Ratio may be 2.5 while for those using it less often it might have been 1.3. This is a sign of a dose-response of the estimated effect of exposure on the outcome.

Incident/incidence refers to newly diagnosed cases; while prevalent cases are any existing cases at any point in time or over a certain period in time. The advantage of using an incident over a prevalent case is that incident cases are not affected by mortality in the same manner that prevalent cases might be. Specifically, if we use prevalent cases we might be mixing up risk factors for disease onset with risk factors for mortality with the disease. PD is a chronic disease with an incipient onset and relatively long duration. Thus, in a PD study with a long prodromal phase prior to diagnosis, it is hard to tell whether a case is truly incident as this might also depend on specialty health care access. Also, it has been suggested that the diagnosis of PD becomes more well established after the first 5 years of onset of motor symptoms, as this period allows movement disorder experts time to distinguish between idiopathic PD and Parkinsonism such as multiple system atrophy or supranuclear palsy.

Confounding is a bias that occurs because a risk factor for the outcome is also a cause or precursor of the exposure of interest, raising the possibility that the outcome is caused by this confounder and not by the exposure that one is trying to evaluate. For example, if gender is a risk factor for PD and gender is also associated with occupational exposure to pesticides, we would want to adjust all effect estimates for pesticides by gender to remove potential confounding bias.

Recall bias is a type of exposure misclassification also known as information bias (see below). This bias is considered 'differential' by epidemiologists, meaning that the exposure is differently misclassified or mis-measured in cases versus controls. For example, cases may remember or report past exposures differently from controls because they have the disease and an incentive to recall or deny exposures. Generally, it has been suggested that cases may put more effort into recalling exposures since they have a need to explain their disease or are more motivated to help researchers while controls are less motivated to recall past exposures or may just not make efforts to look up records etc. Alternatively, if the outcomes of interest are related to cognitive defects it might be the cases who remember less well. The former is most likely a problem if the diseased subject knows or suspects an agent to cause their disease. If the subject has no way to know which pesticide might have caused PD, for example, and is asked to report all chemicals they have ever used occupationally, it is unlikely that they would only recall one but not other chemicals differentially. Thus, if recall bias existed, we would expect that all pesticides they reported to the researchers would show an association with the outcome and not just one amongst many, since the tendency of cases to recall better or more exposures than controls would not be expected to be specific to just one chemical. In fact, when recall has been compared with record-based exposure assessments, differential recall that would cause recall bias has generally not been shown to be a problem. *Note:* non-differential recall error such that both cases and controls misreport their exposures is known to cause mainly a "bias towards the null," i.e., this bias is masking any true effect rather that enhancing them and such random errors in exposure measurement are much more likely to occur than 'recall biases'.

Other biases include information bias which is characterized not only as mismeasurement of exposures but also of outcomes which can severely distort results in both case-control and cohort studies. As long as any mismeasurement is non-differential (see above) i.e. the same for cases and controls or for the exposed and unexposed (in case of disease misclassification), such biases most often cause underestimation of true effect sizes, i.e., the bias results in moving effect estimates towards the null and this underestimation of true effects can be severe. Finally, there is selection-bias if controls are not representative of the exposures in the population that gave rise to the cases in case-control studies, or when there is a large and differential (with regard to case status) loss to follow-up in cohort studies.

2.1 Literature search

To obtain all published studies on the relationship between PD and paraquat, I undertook a literature search using the same method to search for articles that I normally use in my research. This is the same method that I teach my UCLA students to use. As such, I relied upon two search engines, PubMed (https://www.ncbi.nlm.nih.gov/pubmed) and Google Scholar (https://scholar.google.com/). PubMed is an excellent resource for finding papers on the exact topic one is interested in, but it does not do as well in finding papers which were largely about a different topic but may have also briefly reported on the topic of interest. Google Scholar does well in capturing every possible paper of interest, but will often provide many articles not relevant to the subject matter at hand. I used both search engines to be as thorough as possible, but also to identify the most relevant articles. These searches initially yielded 493 articles in PubMed and 6,000+ articles in Google Scholar for epidemiological studies; and over 20,000 articles for animal and mechanistic literature. [Most citations were not immediately relevant to the present question, due to their focus on topics such as animal, cell and other types of basic science studies and mechanistically focused or review papers.]

As is typical in most published reviews or meta-analyses, I took additional steps to ensure I did not miss any relevant articles by also reviewing published papers' citations. For these, I relied on articles on the topic that were published most recently, i.e., after 2016.

2.2 Reliance on peer-reviewed literature

As I teach my students, the most relevant articles, and indeed the only articles I generally review and cite in my own research, are those that have gone through peer review at a reputable journal. In each field there are different journals considered most reputable; but in general, a reputable journal is listed in a well-known and respected indexing source such as PubMed.[i] Typically, such journals have been published for years and many are the official journals of and are backed by well-recognized and respected medical or research non-profit organizations, such

[i] PubMed is a service of the US National Institutes of Health (NIH). On their website (https://www.nlm.nih.gov/pubs/factsheets/j_sel_faq.html) they explain that NIH uses a committee, the Literature Selection Technical Review Committee, to review and recommend which biomedical and health- related life science journals are included. Criteria include relevant subject matter as well as journals that meet PubMed Central's scientific quality standard, described as "scientific and editorial character and quality of a journal."

as the American Medical Association, the British Medical Association, the International Society for Environmental Epidemiology, the American Association of Neurology, or the Movement Disorder Society.

Peer review has been defined as "a system by which manuscripts submitted for publication are evaluated, using outside referees (peers), who comment on the manuscripts' merit, originality, significance, and appropriateness to the journal. The intent is to identify flaws in design and analysis or interpretation, to suggest improvements, to direct manuscripts to the most appropriate outlets, to discourage repetition in publishing, and to weed out poor science or scholarship." [1]

Independent peer review is the cornerstone of science internationally and in the United States and forms the basis for what the scientific community considers acceptable and reliable medical and scientific research. The peer review process is mostly done anonymously, that is the reviewer is anonymous, although this has been changing recently, while generally - but not always - the authors are identified. This system is thought to provide the intellectual rigor required to ensure that scientific research is conducted, and papers written in adherence with the principles acceptable in the field. These principles refer to the epidemiologic and statistical methods employed, such that the research protocols used are considered widely accepted and the methods are rigorous, and the reported results may be deemed valid as they either have avoided or accounted for or assessed the extent of potential biases. Finally, reviewers are asked to assess not only the methods but also the authors' conclusions in the context of the study's findings. Peer reviewers provide their opinions on whether or not an article is acceptable for publication and make recommendations to the editors. As part of the process, peer reviewers assemble comments for the editors who will communicate these and their own concerns to the authors for clarification and request additional information with the intention to improve the manuscript for the readership. Importantly, this process also allows experts in the field to assess the validity of the research (this has become especially important as lengths limitations do not allow authors to describe all of their methods in great detail; with the growth of online supplementary materials this issues is becoming less acute). When any validity issues spotted during the review process that are not addressed sufficiently by the authors in their responses and/or a revised manuscript, the editor may reject the manuscript as not being ready for publication.

To succeed within this system, authors typically will only submit their best work; i.e., work they have rigorously discussed with all co-authors and deemed of high quality after extended discussions. Secondly, authors have to answer to reviewers' critiques and must be willing to make changes requested or need to argue against suggested changes if there is a compelling scientific reason or provide justifications to the editors for not adhering to reviewers' suggestions. The ultimate decision is made by the editor on whether a manuscript is accepted for publication. I have personally peer reviewed on hundreds of occasions and for more than 30 different journals. I have also served on the editorial boards of three journals: Epidemiology, Epidemiologic Perspectives and Innovations, and Environmental Health.

This system of peer review has been practiced for decades. While it is not without fail, most revisions that are suggested by reviewers will improve the quality of the published manuscript; furthermore, this system can head off potential scientific fraud and encourages scientific honesty and the publication of state-of-the-sciences papers [1].

2.3 Conflicts of interest.

Several systematic reviews have been published on the role of conflicts of interest in medical research. For example, a review of 1,140 original studies reported a strong relationship between industry sponsorship and pro-industry conclusions, with it being more than 3 times as likely to find conclusions sympathetic to industry in industry-sponsored studies [pooled Odds Ratio (OR): 3.60, 95% Confidence Interval (CI), 2.63-4.91] [2].

Similarly, a recent (2016) article in the British Medical Journal (BMJ), which analyzed the results of 190 clinical trials published in 2013, reported that a financial tie between study investigators and industry was associated with a more than threefold increase in a study reporting a successful trial result (OR=3.23, 95% CI 1.7-6.1) [3].

These reviews suggest that industry sponsorship and financial incentives are unequivocally related to study findings and this is widely recognized by medical and research communities. This is one reason why journals have become more and more rigorous in requiring extensive reporting of conflicts of interest by investigators for original research and review articles. Potential sources of conflicts are published along with the manuscript and allow the reader to take these into account when drawing conclusions about the validity of the findings and especially of how the authors presented and interpreted their data. This information is also made

available to journal reviewers, as it may influence their reviews in terms of assessing the science presented and also giving them a chance to comment on potential issues with a manuscript before they make a recommendation for publication as the information may contribute to their assessment of the validity of the reported research. Furthermore, this is what I as a professor teach my students, and also what UCLA generally teaches to students in bioethics courses about conflicts of interest.

I performed my own review of the data contained in the literature and provide my opinions on these data throughout this report. There are some studies where a conflict of interest is apparent or possible, i.e., manuscripts published that include authors who have been receiving reimbursement for their work by industry or industry sponsored groups. Even though the authors disclosed such conflicts, there is a likelihood that the funding sources or financial ties to industry may have influenced the science reported and I decided to put less weight on the conclusions of these papers due to a possible lack of impartiality by the researchers.

2.4 Statistical significance.

When we conduct statistical analyses, we are starting off with an assumption of no association between paraquat and Parkinson's disease as the "null hypothesis". This allows us to determine the p-value for this null hypothesis with regards to our findings. Again, the p-value is the probability of obtaining an estimate (of effects) at least as far from a pre-specified value (the estimates' null value in case of the null hypothesis) as the estimate we have obtained, if that specified value were the true value. Recall that a p-value *does not* give the probability that a specified hypothesis is true. As a convention a $p<0.05$ is often considered "statistically significant". It is important to realize that this is simply a convention which can and often is replaced by other p-values such as $p<0.01$ or $p<10^{-7}$ (in genomic studies). The meaning of a p-value of 0.05, given the null hypothesis is true, can be expressed as the following: if you repeatedly conduct 100 tests of samples drawn from the same population (people), then in 5% of your tests, you would obtain the results you got solely due to random error (chance). Thus, it is a metric intended to show the likelihood of random error. It *should not* be interpreted as the probability that paraquat causes or doesn't cause Parkinson's disease. Moreover, if $p>0.05$, this doesn't "prove" that the null hypothesis of no effect (or no causal connection) is true since absence of proof is not proof of absence.

Similarly, when a (95%) confidence interval excludes 1.0 (such as an OR=2.0, 95% CI=1.2-2.8) – because 1.0 (the null value of an OR) is not inside (or covered) by the confidence interval-- it would be considered "statistically significant". As with p-values, confidence intervals can be defined differently as 95% intervals or also as 90% or 80% etc. intervals. However, confidence intervals provide additional information that a p-value does not provide. This is information about the precision of the estimate - sometimes called the 'informativeness' of the data. In practice, p-values and confidence intervals that are close to the null value of the estimate (for example, if one side of the confidence interval is between 0.9 to 1.1) are sometimes considered 'marginal' in terms of statistical significance. It is however important to realize that estimates that are the least influenced by chance are not those with low p-values, but those with narrow confidence intervals. Also, statistical significance does not reflect clinical or other types of significance of a result. Similar to what was described above for the OR or RR, a highly 'statistically significant' result that reflects an increase in children born with a skin mole might not be clinically or public health relevant or for example an increase in 1mm Hg in systolic blood pressure might be statistically significant in a large population but is clinically not actionable or meaningful in an individual.

Statistical significance testing, while widely used for decision making, has often been misused in the medical literature, and therefore it has been widely criticized. One journal now bans the use of all statistical tests and even confidence intervals [4]. Considerable debates on the merits and problems of significance testing have been published. [5] In many Schools of Medicine and Public Health such as UCLA, students have been taught for decades to not simply rely upon statistical significance testing to draw conclusions. This is in accordance with the writings of the faculty member Dr. Sander Greenland, an author of the most widely used textbook in Epidemiology Methods entitled "Modern Epidemiology" [6]. In fact, Dr. Greenland recently co-authored an article published in *Nature* and co-signed by more than 800 scientists including epidemiologists and biostatisticians that again asks for a more nuanced interpretation of statistics and data that is not reduced to statistical significance testing [7]. At UCLA, we teach students to focus on the point estimate (e.g. the Odds Ratio or Rate Ratio) as a measure of the size of the association between exposure and disease and the confidence interval to gauge the precision of this estimate and the informativeness of the data/study. We also teach them that this

only evaluates the random error of the estimates and that any systematic biases need to be assessed through other means that allow us to assess biases (see below).

To evaluate precision and the role that random error plays in data, it is also important to consider the rarity of a disease, because the rarer a disease, the harder it is for a scientist to create a large enough study with enough cases enrolled to generate adequate statistical power. Parkinson's disease (PD) is considered a rare disease. The annual incidence rate (number of new cases) of PD is somewhere between 10-20 cases per 100,000 people [8] [9]. Thus, it is hard to study PD using a cohort study design, because a cohort would have to follow hundreds of thousands of people for many years into older adulthood in order to be able to find a result that would give us a $p<0.05$ if we assume that the effect estimate size is moderate (less than 2). This is the main reason why most rare disease (including PD) studies employ a case-control design which is much more efficient in terms of the sample size necessary to have enough statistical power and also in terms of reducing study costs and increasing feasibility.

Many of the case-control studies cited below in this review, particularly those that tried to recruit cases in rural areas, had a limited sample size simply because there are a finite number of people living and cases to be found in rural areas which have relatively lower population density than metropolitan areas. One of the earliest studies that suggested that pesticide exposures may be responsible for PD was an ecologic study [10] that relied on ICD code and prescription data from the Canadian national healthcare system to identify prevalent PD cases; they found that only about a quarter of the confirmed PD cases lived in rural provinces and that an absence of neurologists in some areas may have been contributing to underdiagnoses. Thus, two Canadian population-based case control studies [11] [12] that focused on rural areas and used a very labor intensive and involved approach for case ascertainment and validation as well as for exposure assessment, were only able to enroll a total of 57 or 127 incident PD patients. In the first study, paraquat was reportedly used by only 4 out of 57 cases and 0 out of 121 controls (one has to replace the zero by 0.5 in order to be able to estimate an OR of 16.9) [11], while in their second study, paraquat exposure was reported by 5 PD patients and 4 population controls (OR=1.25; 95% CI 0.34 - 4.63) [12]; as a result, this study has very wide confidence intervals as the number of cases is small and in addition the exposures are uncommon. Similarly, another early Canadian study [13] provides very limited overall evidence for paraquat as only one case reported paraquat

use. However, interestingly, this exposed subject was the only PD case with an age at onset of PD before 40 years of age.

While we recognize that wide confidence intervals are not unusual in epidemiologic studies of rare diseases like PD, scientists are nonetheless encouraged to move forward and publish their results because small studies can later be summarized in pooled or meta-analyses, which generally improve statistical power to estimate more precise summary effect estimates.

In addition, as is widely accepted and as we also teach at UCLA, one study alone is never definitive. It is important to look at the information in the literature as a whole to understand relationships between exposure and disease. We teach students to consider point estimates (Odds Ratios) as indicators of associations and effect sizes, and to not dismiss or mis-interpret studies that have wide confidence intervals even if they include the null value, but to put all results into the context of other studies.

2.5 Abstracts vs. full articles.

Whenever possible it is preferable to examine and cite a full article over an abstract of the same study, because full articles have the space to provide a detailed overview of study methods and findings. If the full article is not yet published, however, it is common practice to cite abstracts, especially when they have been peer-reviewed for conference presentations.

3. Literature Review.

In the following, I summarize the findings of the epidemiologic studies on paraquat and Parkinson's disease.

In reviewing the literature, the sample sizes and especially the number of cases should be noted, because - apart from exposure prevalence - they bear the most influence on 'statistical significance' tests and the width of confidence intervals. Smaller studies still may convey some important information, even though their findings may only be statistically 'suggestive' due to sample size (and thus wide confidence intervals), and they may also contribute to results summarized in pooled and meta-analyses. The latest meta-analysis by Tangamornsuksan et al., 2018, [14] reviewed the available epidemiology literature on paraquat and PD up to May 2018; for this review the authors selected publications of 13 case-control and one cohort study. Only

studies with mutually exclusive populations – generally selecting the one with the most subjects or best exposure assessment for the chemical of interest - were eligible and selected for inclusion, as is a customary and valid approach. There was only one cohort study published. It is a preferred approach in meta-analyses to summarize results from case-control and cohort studies separately, and accordingly this meta-analysis only used the case-control studies to generate its final meta-analytical summary estimates. The studies included originated from the USA, Canada, Germany, Finland, France, the Netherlands and Taiwan and the meta-analytic summary OR estimate for paraquat exposure for all case-control studies together was based on a total of 3,231 PD cases and 4,901 controls; an overall OR of 1.64 (95% CI 1.27-2.13; measure of heterogeneity I^2=24.8% p=0.19) was reported. These authors furthermore conducted several sensitivity (subgroup) analyses that are informative; for example, according to an author-generated quality score for each study, the design of the case-control study (population, hospital based), and the type (occupational and/or environmental) of exposure assessed; as well as the type of PD cases (prevalent or incident). For these subgroup analyses, they reported ORs ranging between 1.22 and 2.09 depending on how they grouped studies and all but two 95% CIs excluded the null value of 1 (meaning these meta-analytic estimates were formally statistically significant). Generally, these sensitivity/subgroup analyses did not distinguish results sufficiently by these factors, but of interest is that a larger effect estimate size was estimated for a summary of studies that assessed occupational and environmental exposures (OR of 2.02; 95% CI 0.99 - 4.10).

Another earlier meta-analysis report published by Pezzoli et al., 2018, [15] categorized studies according the Newcastle-Ottawa Scale (NOS) (with 9 points representing the highest quality study) and reported an overall effect estimate for PD with exposure to paraquat as 2.19 (95% CI 1.48 – 3.26) based on seven studies. After restricting to the five higher quality studies (NOS 7+) the estimate was 1.72 (95% CI 1.28–2.32), and furthermore this estimate was judged to not be affected by publication bias. The highest quality among these studies (a NOS of 8) was the study by Wang et al., 2011, [16] in the California study that we describe below.

As statistical power also depends on the prevalence, strength, and variation in the exposure of interest when we consider exposure-disease associations, an informative study is a study that has a reasonable exposure prevalence or provides a strong exposure contrast and is affected by as little exposure misclassification as possible. Thus, some of the strongest studies are those that either assess occupational exposures or investigate highly exposed agricultural

populations and use well-documented and accepted exposure measures. In the following, I will first describe studies with strong exposure assessment or contrasts and then summarize results from other epidemiologic studies that are more limited in their contributions.

A case-control study from Taiwan reported on residential and occupational paraquat exposures, i.e., specifically farming related exposures and PD [17]. These researchers assembled 120 patients with PD from the Movement Disorder Clinic of National Taiwan University Hospital in Taipei between July 1993 and June 1995. Therefore, the diagnostic validity of a PD diagnosis is high. In this hospital-based study, they matched 240 hospital control subjects to the PD patients by age (mean age ~63 years) and gender; these controls were recruited from the neurologic or medical outpatient clinics at the same hospital serving the same geographic population. The controls were from outpatient clinics that treated headaches, back pain, cervical spondylosis, and peripheral neuropathy and they excluded any patients who might have had exposure to neuroleptics and a previous diagnosis of diseases such as stroke, brain infections, and dementia. All controls were also examined by study neurologists to ensure that they did not have PD, thus further reducing the likelihood for misclassification of the outcome. Information from all subjects was collected according to a structured open-ended questionnaire by trained interviewers in a face-to-face interview with subjects and their family members, if available. Subjects were asked to identify residential and occupational exposures to herbicides and pesticides, chemicals, heavy metals, and minerals. Forty PD patients and control subjects were re-interviewed 4 to 10 months later and all data provided at the re-interview were found to be identical to the initial interviews, i.e., confirming the reliability of the exposure information collected. Increased PD risks were observed for rural living, rice but not orchard farming and occupational or residential exposure to herbicides/pesticides. The strongest (and statistically significant) increase in PD risk was observed for reported use of paraquat (OR=3.2, 95% CI=2.41-4.31). Also, in uni- as well as multivariate conditional logistic regression models, the occupational use of paraquat for more than 20 years was even more strongly associated with PD risk OR=6.41 (95% CI = 2.74-15.00). These authors also reported that the OR for PD in study subjects who previously used both paraquat and other herbicides/pesticides was 4.74-fold higher compared with subjects who used herbicides/ pesticides other than paraquat only (i.e., essentially comparing paraquat exposed to other pesticide exposed subjects). This study is of high quality in terms of its diagnostic validity and exposure reporting reliability – of note: paraquat was mainly

used by rice farmers and most of the increase in risk was seen in this type of farming consistent with the reported exposure to paraquat. Also, long-term exposures resulted in the largest risk increase. Long term use is less affected by non-differential exposure misclassification due to random error in reporting, as this type of use is likely to be well remembered and reported.

The largest cohort study of pesticide use in the US generally is the Agricultural Health Study (AHS) that enrolled licensed private pesticide applicators – mainly farmers - and their spouses in a prospective cohort from 1993 to 1997 in Iowa and North Carolina (n = 84,739). The AHS collected lifelong pesticide exposure information among its cohort members at enrollment (1993–1997) and self-reported physician diagnoses of PD at baseline (78 incident cases) or after 5 years of follow-up in 1999–2003 (83 prevalence cases); the researchers additionally reviewed state mortality records for causes of death. Comparing cases with all cohort members who did not report PD (79,557 at enrollment, 55,931 at follow-up), the researchers reported that cumulative days of pesticide use at enrollment and personally applying pesticides increased the risk of incident PD [18]. Results for 'ever use' of any specific pesticide and PD were mostly null. For paraquat ever use at baseline and prevalent PD they estimated an OR of 1.8 (95% CI 1.0, 3.4), but for incident PD (collected at follow-up) the estimate was null (OR=1.0; 95% CI 0.5, 1.9) suggesting no increase in risk. However, this null result may be due to a large number of cohort members not responding to the PD question (N=28,621) at follow-up; also, loss to follow-up was relatively extensive as only 68% of all cohort members could be reached again, with some of the lost subjects having died or been too ill to participate again. Therefore, it is possible that PD patients who were ill might not have been reached during follow-up. Interestingly in the AHS, paraquat was one of the least ('ever') used herbicides at baseline.

A much more extensive effort to validly assess pesticide use and PD in the AHS study was undertaken in a nested case-control study [19] (The Farming and Movement Evaluation study or "FAME") that took care to 1) clinically confirm the PD diagnosis and 2) achieve much more accurate exposure assessment for pesticides via retrospective interviews. In the FAME study, a total of 115 PD cases (83% of all suspected incident cases) and 383 (71%) eligible controls participated. Study staff (among them neurologists) clinically assessed suspected case subjects at their homes, collecting neurological history, and videotaping a neurologic examination they conducted in a standardized fashion. The final PD diagnosis was then

determined as the consensus of two movement disorder specialists who based it on all available information, including medical records and the videotaped examination, while applying widely accepted diagnostic criteria (National Institute of Neurological Disorders and Stroke/UK Brain Bank criteria). Controls were identified by stratified random sampling from amongst all living and nondemented AHS participants without PD who were also frequency matched to cases by age, gender, and state (Iowa or North Carolina) at a ratio of three controls per case. Trained interviewers used structured computer-assisted telephone interviews (CATIs) to collect complete lifetime occupational histories (including all farm jobs after age 14 and detailed information on pesticide use) from all FAME subjects. The study allowed proxy informants to answer for those subjects unable to complete interviews because of death, hearing or speech deficits, or cognitive impairment. Pesticide exposure measures were generated according to a reference age, defined as age at diagnosis for cases, and as median case diagnosis age in the corresponding gender-, state-, and age-specific stratum for controls. In FAME the researchers determined whether subjects ever used paraquat (mixed or applied one or more times) but also calculated cumulative lifetime years of use after 1965, when paraquat was first marketed in the United States. This study reported a very strong OR for PD with ever being occupationally exposed to paraquat (OR=2.5, 95% CI=1.4-4.7). In a subset of the FAME study - 87 PD subjects and 343 matched controls - the researchers examined gene-environment interactions to investigate the hypothesis that glutathione transferases provide some cellular protection against oxidative stress and therefore may modulate paraquat toxicity [20]. They found that PD risk was higher with paraquat use in individuals with homozygous deletions of the genes encoding glutathione S-transferase T1 (GSTT1) i.e., in men with functional deletions, the estimated effect was extremely strong (OR of 11.1; 95% CI: 3.0 – 44.6; P for interaction of 0.027).

The effect estimates for PD and paraquat from the AHS FAME study are greater than the overall estimates we derived in our own California study (see below) for ambient residential and/or workplace exposures, and are likely indicative of differences in the types of exposure, i.e. occupational exposure due to active handling of paraquat versus ambient or by-stander exposures from paraquat applications near workplaces and residences.

Based on the studies above - of which the Tanner AHS study received the highest quality rating score of 9 on the NOS scale in the review by Pezzoli et al., 2013 [15] - occupational

exposure to paraquat, that involves direct exposure from spraying, mixing and loading, thus increases the risk of developing PD between 2.5 and 6.0 fold depending on the number of years the person applied paraquat.

My own group's analyses for paraquat and PD [16] [21] [22] are based on a long-running population-based case-control study of PD conducted in California's agricultural central valley. This study is the only case-control study to date that did not rely on self-reported paraquat use, as it employed the California pesticide use reporting system records. Specifically, we conducted a population-based case-control study that enrolled patients with new onset (incident) idiopathic PD in two waves from 2001 through 2015 and also population controls from 2002 through 2011 in three counties in central California (Kern, Fresno, and Tulare) from among county residents age 35 years or older who had lived in California for 5 or more years prior to recruitment. PD patients were first invited through neurologists, large medical groups, and public service announcements and later through the CA PD registry pilot program for those diagnosed after 2007. Initially, population-based controls older than 64 years were identified from Medicare lists and younger controls from residential tax assessor records. After the instatement of the Health Insurance Portability and Accountability Act (HIPAA), all controls were randomly selected from residential tax assessor records. The study applied two different sampling strategies for enrollment of controls: 1) mailings to randomly selected residential parcels (or housing units) and follow-up through phone; and 2) random selection of clusters of five neighboring households for staff to conduct home visits and enroll one control per household in person at the doorstep.

Ambient residential and workplace pesticide exposures were assessed using address histories and a geographic information system (GIS)-based model that was validated with serum biomarkers for long-half life organochlorine pesticides [23] and a methylomics biomarker for short half-life organophosphate pesticides [24]. Pesticide use report data collected since 1974 by the California Department of Pesticide Regulation, computerized historical land use on maps from the California Department of Water Resources, and geocoded address information were combined in this GIS model to generate annual exposure estimates for each active chemical ingredient of (restricted use only until 1989, all thereafter) pesticides used in California. Pesticide-use reports (PURs) are collected for all commercial applications of pesticides including agricultural applications. Each PUR record includes the name of the pesticide's active ingredient, the poundage applied, the crop and acreage of the field, the application method, the date of

application, and a PUR locator, which can be linked to the Public Land Survey System (PLSS), a nationwide grid that parcels land into sections of approximately one square mile (640 acres). The GIS and geocoding approach has been described more extensively in the peer-reviewed literature [25] [26]. We used it to estimate at home and workplace specific ambient pesticide exposures from applications to agricultural crops for each pesticide active ingredient reported from 1974 onward to time of diagnosis for cases and similarly for controls. Specifically, we estimated pounds of a specific pesticide that were applied within a 500-meter buffer of a residential or workplace address and weighed the total poundage reported by the proportion of acreage treated. For each pesticide we then created period-specific averages by summing over the years and dividing by lengths of the study period prior to PD diagnosis or index date for controls. Participants who at some point in this period did not work or live in California were considered unexposed.

The manuscript with the largest number of Parkinson Environment Gene Study (PEG) cases and controls we published is Lee et al. (2012); it included 357 PD case and 754 control subjects. As described above, we generated a measure of average pounds of paraquat applied per acre within a 500-m buffer around the each home or workplace address, in the study period for each participant's residential and workplace address from 1974 to 1999 (summing over the year specific values and dividing by the lengths of the total study period). We considered participants ever exposed to paraquat if they had an average study period exposure greater than 0 at both their residential and their workplace addresses. The adjusted OR for ambient ever/never paraquat exposures at the residential and occupational addresses was 1.36 (95%CI 1.02–1.81). Costello et al. (2009 [21]) reported additional subgroup analyses for 368 incident PD cases and 341 population controls with residential exposures only (ignoring work place address exposures): during the whole period (1974-1999), participants who were of younger age (less than 60 years) at the time of diagnosis were at risk when ever-exposed to either paraquat or maneb alone (OR of 1.77; 95% CI 0.84, 3.75) and much increased risk when exposed to both pesticides residentially (OR of 5.07; 95% CI 1.75, 14.71). While ever paraquat exposure alone near the residence did not increase risk (OR of 1.01; 95% CI 0.71, 1.43) among all PD patients, the combined exposure at residences increased risk quite strongly (OR of 1.75; 95% CI 1.13, 2.73). These differences might be due to types of applications or intensity of use being different on crops that were treated with just one or both pesticides, as well as synergisms in exposure effects on health or exposure

misclassification bias towards the null for less intensively treated fields. In Wang et al. (2011 [16]), we explored combinations of ambient exposure to three pesticides (paraquat, maneb and ziram) and also estimated effects at workplaces (OR: 3.09; 95% CI: 1.69, 5.64) as well as at residences (OR: 1.86; 95% CI: 1.09, 3.18). Additional adjustment for exposure to organophosphate and organochlorine pesticides shifted the risk estimates only slightly but – as expected - widened confidence intervals; yet, the combined exposure to maneb, ziram, and paraquat at workplaces remained strongly associated with PD risk even after adjustment for organophosphates and organochlorines (OR: 2.61; 95% CI: 1.24, 5.48). Overall, ambient paraquat exposure alone, at both residences and workplaces, increased the risk with an OR of 1.50 (95% CI 1.03, 2.18).

Another clinic based case-control study recruited 519 PD cases from eight North American movement disorder centers between July 1, 2004, and May 31, 2007 [27]. They evaluated diagnostic features for parkinsonism and also enrolled 511 controls who were frequency matched to cases by age, sex, and location. They identified these controls from (1) nonblood relatives (excluding spouses) or acquaintances of cases enrolled in the study or (2) nonblood relatives or acquaintances of other patients with parkinsonism treated in the same movement disorders clinics but excluded persons with neurodegenerative disorders as well as workmates. Additionally, some male controls aged 50 to 65 years were recruited through commercial telephone lists within zip codes matching those of enrolled cases.
Exposure was assessed in standardized computer-assisted telephone interviews, including collecting detailed information for the following job tasks using pesticides. Only 144 (16%) of the study population reported that they had farmed at some point in their lives. Overall there were 71 pesticide users, 51 (72%) farmers and 18 who had worked in building and grounds maintenance and 2 soldiers. While paraquat was rarely used (9 cases, 4 controls), this study still suggested positive and strong association with a risk of developing PD (OR=2.80 95% CI (0.81-9.72).

In the French agricultural cohort AGRICAN [28], researchers assessed associations cross-sectionally between self-reported prevalent PD (N=1,732 cases) and pesticide use derived from information on types of livestock on a farm and crops grown. Specifically, these authors collected a history of lifetime exposure to 13 crops and 5 types of animals and some self-reported pesticide use at enrollment (2005–07) among 181,842 study participants and generated a crop-

exposure matrix (PESTIMAT) to estimate exposure to selected active ingredients and duration of use. Specifically, for a given crop and active ingredient, annual exposure was based on a combination of information on pesticide registration, sales and recommended use. Study participants were considered exposed to an active ingredient if: (i) they stated that they had cultivated the crop; (ii) reported personally treating the crop with a pesticide; and (iii) the active ingredient was registered and recommended for use on the crop during that year. Associations between pesticide use and PD were estimated by logistic regression according to crops and livestock, adjusting for potential confounding factors (sex, age, educational level, smoking status, alcohol consumption). Paraquat use was associated with an increased risk of PD in single pesticide models (OR= 1.43; 95% CI 1.17, 1.75); when these researchers adjusted for co-exposures between active ingredients, all associations except for two low use pesticides moved towards the null of showing no association and for paraquat the OR became 1.01 (95%CI 0.41, 2.49). However, these adjusted models are likely over-adjusting for highly correlated pesticide exposures (due to the use of a crop exposure matrix) with a resulting 'multi-collinearity' that is expected to bias estimates towards the null for the most highly used and correlated pesticides such as paraquat. The authors concluded that their study did provide additional evidence for an association of paraquat with PD.

Finally, there are some smaller studies and some less rigorously conducted studies in terms of design or exposure assessment, a few have already have been mentioned above. Briefly, an early Canadian study was an ecologic [10] and identified prevalent PD cases based on ICD code and prescription data from the Canadian national healthcare system and was amongst the first to suggest that pesticides used in agriculture may contribute to PD risk. A very recent study conducted in the US state of Louisiana also employed an ecologic design mapping residential zipcodes for 23,224 primary PD discharge diagnoses between 1999 and 2012 against known crop distributions [29]. This study suggested associations for major PD-affected areas of the state with arbor-pastoral pesticides in general but also specifically listed paraquat as one of those chemicals. Even though this Louisiana study identified a large number of PD cases, it did not distinguish between urban and rural residents, had to rely on a single residential address of prevalent PD cases, and also had no information on individual level pesticide exposures. Furthermore, there is the possibility of the results being affected by the well-known ecologic study bias.

A study by Engel et al. (2001 [30]) neurologically examined a cohort of 310 subjects who had been occupationally exposed to pesticides. However, this study was only able to generate a 'Parkinsonism' measure based on the presence of common and non-specific motor symptoms rather than studying the much rarer outcome of idiopathic PD. The OR calculated for the association of paraquat with the Parkinsonism symptom complex was null (adjusted OR = 0.8; 95% CI 0.5 to 1.3). Firestone et al. (2005 [31]), conducted a population-based case-control study in the Group Health Cooperative health care system in western Washington State and included 156 incident PD case patients and 241 healthy age- and sex matched control subjects with occupational self-reported pesticide exposures derived from a structured interview in their analyses. They reported an elevated OR for PD and paraquat (OR, 1.67; 95% CI, 0.22-12.76), but this estimate was extremely imprecise as it was based only on 2 cases and 2 controls who reported having used paraquat occupationally. Dhillon et al. (2008 [32]) used a clinic-based case-control design to study PD and self-reported exposure to pesticide products, specifically focusing on organic pesticides such as rotenone, but also collected information on other occupational and environmental exposures. All PD patients and controls were from the same neurologic clinic in East Texas but controls had no history of PD. A total of 4 cases and 1 control reported ever personally mixing or applying paraquat and the OR was reported as 3.5 (95% CI 0.4–31.6). Rugbjerg et al. (2011 [33]) conducted a population-based case–control study in British Columbia, Canada, identifying PD cases from reimbursements for anti-parkinsonian agents and selecting controls (frequency-matched to the case sample by birth year, gender, and geographic region) from the universal health insurance database. A total of 403 cases and 405 controls were interviewed about their job, medical and personal habits histories, and beliefs about disease risk factors. Among those reporting pesticide exposure, an occupational hygiene review of jobs was conducted and out of 121 persons who self-reported pesticide exposures, 53 were excluded because the reported exposure was judged to be limited (e.g. sales personnel handling closed containers or construction workers occasionally handling wood treated with preservatives) while those remaining were mainly farmers and forestry personnel. In the end, the authors of this paper reported that only three cases and three controls had reported paraquat exposure and they did not provide any effect estimates for their study for paraquat. Three additional early Canadian studies have already been mentioned above. Briefly, in Hertzmann et al. (1990 [11]), paraquat was reportedly used by only 4 out of 57 cases and 0 out of 121 controls (one has to replace the zero

by 0.5 in order to be able to estimate an OR of 16.9). In Hertzman et al. (1994 [12]), paraquat exposure was reported by 5 PD patients and 4 population controls (OR=1.25; 95% CI 0.34 - 4.63). And finally, Semchuk et al. (1992 [13]) reported that only one case used paraquat, and this person was the only PD case with an age at onset of PD before 40 years of age.

Caballero et al. (2018 [34]) conducted a recent mortality record based study in Washington State, specifically, they identified PD as a cause of death between 2011–2015 and geocoded the residential addresses on the death certificates in order to classify them according to exposure to agricultural land-use within 1000 meters. Individuals exposed to land-use associated with paraquat application were increased (OR = 1.22, 95% CI = 0.99–1.51). Mortality records, however, severely underestimate PD incidence and prevalence, as not all PD affected subjects are recorded as having suffered from PD as the underlying or even contributing cause of death, but the effect estimates are consistent with those of the higher quality studies.

Industry-sponsored studies

A meta-analysis of epidemiologic studies (most of which have been discussed above) was sponsored by Syngenta and reported results very similar to the above cited meta-analysis by Tangamornsuksan et al. (2018 [14]), i.e. a similar size increase in PD with any type of paraquat exposure (random effects meta-OR: 1.47, 95% CI 1.01-2.13) as reported in what the authors called tier 1 and 2 studies (fixed effects model for all studies RR = 1.69 (95% CI = 1.44–1.98). The data pooled in the Breckenridge et al. (2016 [35]) meta-analysis also supported a possible dose response with exposure intensity in a subgroup analysis where this group summarized results for 'very high' paraquat exposure and estimated a random-effects meta-analytic OR of 1.99 (95% CI 0.84-4.71) (fixed effects meta-analysis RR = 1.75; 95% CI = 1.19–2.57). This generally supports the consistency of findings across different settings and study types.

There were also two industry sponsored worker studies, one that examined 18 workers in a paraquat production plant in the UK (average length of employment 5 years) and 18 Malaysian production workers (average employment 2.3 years) [36], the second study examined Malaysian plantation worked who applied paraquat [37]. The first study (Howard 1979) examined clinical

records, medical and occupational histories, with special attention to the skin of workers evaluating rashes, nail damage and epistaxis as a result of direct contact of skin and mucous membranes with paraquat [36]. This study found no evidence of long-term effects on skin, mucous membranes or general health after exposure to paraquat in these workers. However, this study was extremely small and workers did not work in production for long. Most importantly it is unclear how the workers were selected from amongst all workers, which makes severe healthy worker selection bias very likely. Finally, longer terms neurologic outcomes were not evaluated in these production workers.

The second study focused on pulmonary function (FVC, FEY, FEV,%, and single breath CO2 diffusion), renal function (serum creatinine and BUN), liver function (serum ALT, AST, and ALKP) and performed a haematological screen in 27 paraquat spraymen (mean spraying time 5.3 years) compared with two controls – the first consisted of 24 general plantation workers with minimal exposure to paraquat arising from occasional work in recently sprayed areas, and the second consisted of 23 latex factory workers with no known occupational exposure to paraquat [37]. The authors claimed that there were "no significant differences as a consequence of occupational exposure to paraquat" and that "long-term paraquat spraying at the concentrations used produced no quantifiable harmful effects on health as measured by the indices selected for this study", however, a closer inspection of the results shows that being a 'sprayman' was negatively associated with all of the lung function and blood parameters measured and similar to smoking for the lung function parameters. Again, no neurologic exams were conducted.

Another industry sponsored study relied on published cases of suicidal or accidental paraquat poisoning for individuals who either recovered or lived for at least 30 days (70 cases) or lived for 15 to 30 days after paraquat poisoning (13 cases) to assess the consequences of 'high-dose' paraquat exposure [38]. The authors of this poisoning cases review stated that there were no signs of parkinsonism reported in the literature for any such cases and, thus, claim that "no connection between high-dose paraquat exposure in humans and the development of parkinsonism" exists. However, the authors did not report how long any of these so-called "long-term" survivors lived and how long after the poisoning event they were neurologically evaluated. It is also unclear whether all of these patients' health outcomes were recorded by experts who are able to recognize signs of parkinsonism, or whether neurologic exams had even been conducted.

No mention of parkinsonian symptoms in a published report may simply mean that no neurologic evaluations had been conducted as this review refers to "neuro-evaluable" poisoning patients and not patients who actually were neurologically examined. Furthermore, these patients were quite young – median age of 22 years at time of poisoning – and for PD symptoms that developed later in life to be documented, patients would have had to be followed for a long time; yet, the longest post-poisoning follow-up was 10 years. Most importantly, however, the authors assumed that the most likely mechanism of causing PD would be analogous to MPTP: acute high-dose MPTP exposures cause parkinsonism within days of exposure in the absence of systemic toxicity involving other organ systems. Thus, these authors expected that parkinsonism would occur within a short time after poisoning. They, however, contradicted this assumption in their discussion stating that "paraquat neurotoxicity is distinct from that of MPTP and rotenone" and cite Richardson et al. (2005 [39]). Also, high paraquat doses cause not only strong but acutely lethal systemic toxicity involving other organ systems (mostly lung and kidney) different form MPTP.

Finally, Tomenson and Campbell (2011 [40]) conducted a study that used mortality records to identify PD among UK worker employed in 4 plants that engaged in the manufacture of paraquat between 1961 and 1995. They followed 926 male and 42 female workers through 30 June 2009 and compared the observed and expected PD mortality of males using national and local rates. The average age of male employees at first exposure was 32.8 years, and I calculated that the average follow-up time must have been 31 years. Thus, the average age at last follow-up would have been 64 years of age which is below the median age at PD onset (68 years) in most population-based studies. Because the cohort is very small and the age at last follow-up on average still young for PD, we would expect very few cases. Indeed, amongst the 292 male workers who had died by end of follow-up (30 June 2009), only for one man was the cause of death attributed to PD. The authors report that they would have expected 1.8 cases using national or local rates for PD as an underlying cause of death and they calculated a standardized mortality ratio of 0.55 (95% CI 0.1 to 3.09) for all 926 male workers. However, in this workforce only approximately 300 men were considered exposed to paraquat at medium or high levels and the worker who died of PD had a medium level exposure to PQ. Thus, if we would only consider these 300 workers as truly paraquat exposed workers, based on the data presented in the article, I estimated that - based on local or national rates - we would now expect to find 0.57 cases (about

1/3 of the 1.8 cases we would expect in 926 workers) and in exposed workers this would result in an estimated relative mortality ratio of 1.75; yet, I would expect the 95% CI to be extremely wide as all of these calculations are based on only one observed case and, thus, are statistically highly uncertain rendering this study generally uninformative. Nevertheless, the number of PD cases found in this study cohort is likely an underestimate, as it has been estimated that only 37% to 76% of known PD cases in Britain were recorded on death certificates during this period of time [41].

Bradford-Hill criteria evaluation

The strength (effect size) criterion is partially met since the overall meta-analytical (point) effect estimates reported for ever never paraquat use are about 1.4 reflecting a weak to moderate size association. However, the effect estimates for longer or more extensive exposures in several studies were larger, i.e., close to and above 2, and this can be considered a stronger endorsement of a causal relation; it is further supported by the observed dose response (biological gradient such that risk increases with dose - another Bradford Hill criterion) that some studies found (also note: a small association does not mean that there is not a causal effect, though the larger the association, the more likely that it is unbiased and thus causal). In terms of consistency, this criterion is met since positive associations have been reported for different populations and in different places and different time periods which strengthens the likelihood of a true effect. Temporality, i.e., that PD occurred after exposure and that there is an expected delay between the cause and effect has also been reported, i.e., exposures were assessed and recorded for the periods prior to PD occurrence. Unfortunately, few studies examined the influence of exposure lagging but the ones that did confirmed effects [21]. The specificity criterion (i.e. that one specific exposure causes one specific outcome) is hard to apply in the case of herbicide or pesticide exposure since almost none of the farmers/pesticide applicators or residents near farming operations would be expected to solely be exposed to paraquat, since most farming operations use multiple pesticides over time. Finally, other types of studies (experimental animal and cell studies) suggested certain pathways involved in neurodegeneration in PD increasing biologic plausibility for the action of paraquat on the oxidative stress pathways (see also below).

Biological plausibility.

Biomonitoring studies affirm that some (not all) persons who apply paraquat occupationally have measurable paraquat levels excreted in urine. [42] [43] [44] [45] [46] [47] [48] [49]

While a *plausible mechanism* between cause and effect is helpful, Bradford Hill noted that knowledge of the mechanism is often limited by current knowledge; nevertheless in neuroscience laboratories worldwide scientists have treated C57BL/6 mice with paraquat to induce PD pathology and neurobehavioral effects consistently. As a potent redox cycler causing oxidative stress, paraquat induces two pathologic hallmarks of PD: selective loss of dopaminergic neurons in the substantia nigra pars compacta and an up-regulation of alpha-synuclein (which leads to Lewy body formation in humans) [50]. Neurobehavioral effects including lack of grip strength and postural instability are consistent with the loss of motor function in PD. The same pathological hallmarks of PD were also shown in a long-term rat study [51].

I have been informed that a Syngenta scientist, Louise Marks, also replicated paraquat findings in the independent literature. She found that paraquat treatment in C57BL/6 mice induced the pathological hallmarks of PD in 3 separate studies between 2000-2007. But, the results of her studies were not fully disclosed to the EPA until December 2019. Further, I understand that these studies have never been published.

Finally, while *coherence* between epidemiological and laboratory findings increases the likelihood of a true effect, Bradford Hill noted that "... lack of such [laboratory] evidence cannot nullify the epidemiological effect on associations." Here, however, we have *coherence* between paraquat epidemiology studies and laboratory findings in paraquat treated animals.

4. Conclusions

The epidemiologic studies as a whole support an increased risk of PD with exposure to paraquat. Due to the rarity of this disease, many of the earlier studies were small in size, leading to wide confidence intervals; yet findings were consistent with nearly all studies having point estimates above 1.0. In the pooled and meta-analyses, results are consistent and unequivocal. The few studies that assessed dose generally found that higher levels of exposure were associated

with increased risk. Because of the combined animal data that is consistent with the etiology of PD and the epidemiology studies establishing an increased risk of PD with exposure to paraquat, in my opinion, to a reasonable degree of scientific certainty, paraquat causes PD.

Beate Ritz, M.D., Ph.D.

Date: July 10, 2020

References cited

[1] B. P. Dancik, "Importance of peer review," *The Serials Librarian,* vol. 19, no. 3-4, pp. 91-94, 1991.

[2] J. E. Bekelman, Y. Li and C. P. Gross, "Scope and impact of financial conflicts of interest in biomedical research," *JAMA,* vol. 289, no. 4, pp. 454-465, 2003.

[3] R. Ahn, A. Woodbridge, A. Abraham, S. Saba, D. Korenstein, E. Madden, W. J. Boscardin and S. Keyhani, "Financial ties of principal investigators and randomized controlled outcomes: cross sectional study," *BMJ,* vol. 356, p. i6770, 2017.

[4] D. Trafimow and M. Marks, "Editorial," *Basic and Applied Social Psychology,* vol. 37, no. 1, pp. 1-2, 2015.

[5] Z. ST and McCloskey DN, The cult of statistical significance: How the standard error costs us jobs, justice and lives., University of Michigan Press, 2008.

[6] K. J. Rothman, S. Greenland and T. L. Lash, Modern Epidemiology, 3 ed., Philadelphia: Wolters Kluwer Health/Lippencott Williams & Wilkins, 2008.

[7] V. Amrhein, S. Greenland and B. McShane, "Retire statistical significance," *Nature,* vol. 567, pp. 305-307, 2019.

[8] Z. X. Zhang and G. C. Roman, "Worldwide occurrence of Parkinson's disease: An updated review," *Neuroepidemiology,* vol. 12, no. 4, pp. 195-208, 1993.

[9] S. K. Van Den Eeden, C. M. Tanner, E. L. Bernstein, R. D. Fross, A. Leimpeter, D. A. Bloch and L. M. Nelson, "Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity," *American Journal of Epidemiology,* vol. 157, pp. 1015-1022, 2003.

[10] A. Barbeau, M. Roy, G. Bernier, G. Campanella and S. Paris, "Ecogenetics of Parkinson's disease: prevalence and environmental aspects in rural areas," *Can J Neurol Sci,* vol. 14, pp. 36-41, 1987.

[11] C. Hertzman, M. Wiens, D. Bowering, B. Snow and D. Calne, "Parkinson's disease: a case-control study of occupational and environmental risk factors," *Am J Ind Med,* vol. 17, pp. 349-355, 1990.

[12] C. Hertzman, M. Wiens, B. Snow, S. Kelly and D. Calne, "A case-control study of Parkinson's disease in a horticultural region of British Columbia," *Mov Disord,* vol. 9, no. 1, pp. 69-75, 1994.

[13] K. M. Semchuk, E. J. Love and R. G. Lee, "Parkinson's disease and exposure to agricultural work and pesticide chemicals," *Neurology,* vol. 42, pp. 1328-1335, 1992.

[14] W. Tangamornsuksan, O. Lohitnavy, R. Sruamsiri, N. Chaiakunapruk, C. N. Scholfield, B. Reisfeld and M. Lohitnavy, "Paraquat exposure and Parkinson's disease: a systematic review and meta-analysis," *Arch Env Occup Health,* p. 1492894, 2018.

[15] G. Pezzoli and E. Cereda, "Exposure to pesticides or solvents and risk of Parkinson's disease.," *Neurology,* vol. 80, pp. 2035-2041, 2013.

[16] A. Wang, S. Costello, M. Cockburn, X. Zhang, J. Bronstein and B. Ritz, "Parkinson's disease risk from ambient exposure to pesticides," *Eur J Epidemiol,* vol. 26, pp. 547-555, 2011.

[17] H. H. Liou, M. C. Tsai, C. J. Chen, J. S. Jeng, Y. C. Chang, S. Y. Chen and R. C. Chen, "Environmental risk factors and Parkinson's disease: a case-control study in Taiwan," *Neurology,* vol. 48, pp. 1583-1588, 1997.

[18] F. Kamel, C. M. Tanner, D. M. Umbach, J. A. Hoppin, M. Alavanja, A. Blair, K. Comyns, S. M. Goldman, M. Korell, J. W. Langston, G. W. Ross and D. P. Sandler, "Pesticide exposure and self-reported Parkinson's disease in the Agricultural Health Study," *Am J Epidemiol,* vol. 165, no. 4, pp. 364-374, 2006.

[19] C. M. Tanner, F. Kamel, G. W. Ross, J. A. Hoppin, S. M. Goldman, M. Korell, C. Marras, G. S. Bhudhikanok, M. Kasten, A. R. Chade, K. Comyns, M. B. Richards, C. Meng, B. Priestley, H. H. Fernandez, F. Cambi, D. M. Umbach, A. Blair, D. P. Sandler and J. W. Langston, "Rotenone, paraquat, and Parkinson's disease," *Environ Health Perspect,* vol. 119, pp. 866-872, 2011.

[20] S. M. Goldman, F. Kamel, G. W. Ross, G. S. Bhudhikanok, J. A. Hoppin, M. Korell, C. Marras, C. Meng, D. M. Umbach, M. Kasten, A. R. Chade, K. Comyns, M. B. Richards, D. P. Sandler, A. Blair, J. W. Langston and C. M. Tanner, "Genetic modification of the association of paraquat and Parkinson's disease," *Mov Disord,* vol. 27, no. 13, pp. 1652-1658, 2012.

[21] S. Costello, M. Cockburn, J. Bronstein, X. Zhang and B. Ritz, "Parkinson's disease and residential exposure to maneb and paraquat from agricultural applications in the central valley of California," *Am J Epidemiol,* vol. 169, pp. 919-926, 2009.

[22] P.-C. Lee , Y. Bordelon, J. Bronstein and B. Ritz, "Traumatic brain injury, paraquat exposure, and their relationship to Parkinson's disease," *Neurology,* vol. 79, pp. 2061-2066, 2012.

[23] B. Ritz and S. Costello, "Geographic model and biomarker-derived measures of pesticide exposure and Parkinson's disease," *Ann NY Acad Sci,* vol. 1076, pp. 378-387, 2006.

[24] K. C. Paul, Y.-H. Chuang, M. Cockburn, J. M. Bronstein, S. Horvath and B. Ritz, "Organophosphate pesticide exposure and differential genome-wide DNA methylation," *Science of the Total Environment,* vol. 645, pp. 1135-1143, 2018.

[25] D. W. Goldberg, J. P. Wilson, C. A. Knoblock, B. Ritz and M. G. Cockburn, "An effective and efficient approach for manually improving geocoded data," *Int J Health Geogr,* vol. 7, p. 60, 2008.

[26] R. P. Rull and B. Ritz, "Historical pesticide exposure in California using pesticide use reports and land-use surveys: an assessment of misclassification error and bias," *Environ Health Perspect,* vol. 111, pp. 1582-1589, 2003.

[27] C. M. Tanner, W. Ross, S. A. Jewell, R. A. Hauser, J. Jankovic, S. A. Factor, S. Bressman, A. Deligtisch, C. Marras, K. E. Lyons, G. S. Bhudhikanok, D. F. Roucoux, C. Meng, R. D. Abbott and J. W. Langston, "Occupation and risk of parkinsonism: a multicenter case-control study," *Arch Neurol,* vol. 66, no. 9, pp. 1106-1113, 2009.

[28] C. Pouchieu, C. Piel, C. Carles, A. Gruber, C. Helmer, S. Tual, E. Marcotullio, P. Lebailly and I. Baldi, "Pesticide use in agriculture and Parkinson's disease in the AGRICAN cohort study," *International Journal of Epidemiology,* vol. 47, no. 1, pp. 299-310, 2018.

[29] M. E. Hugh-Jones, R. H. Peele and V. L. Wilson, "Parkinson's disease in Louisiana, 1999-2012: based on hospital discharge diagnoses, incidence, and risk in relation to local agricultural crops, pesticides, and aquifer recharge," *Int J Environ Res Public Health,* vol. 17, p. 1584, 2020.

[30] L. S. Engel, H. Checkoway, M. C. Keifer, N. S. Seixas, W. T. Longstreth Jr, K. C. Scott, K. Hudnell, W. K. Anger and R. Camicioli, "Parkinsonism and occupational exposure to pesticides," *Occup Environ Med,* vol. 58, pp. 582-589, 2001.

[31] J. A. Firestone, T. Smith-Weller, G. Franklin, P. Swanson, W. T. Longstreth and H. Checkoway, "Pesticides and risk of Parkinson's disease: a population-based case-control study," *Arch Neurol,* vol. 62, pp. 91-95, 2005.

[32] A. S. Dhillon, G. L. Tarbutton, J. L. Levin, G. M. Plotkin, L. K. Lowry, J. T. Nalbone and S. Shepherd, "Pesticide/environmental exposures and Parkinson's disease in East Texas," *J Agromed,* vol. 13, no. 1, pp. 37-49, 2008.

[33] K. Rugbjerg, M. A. Harris, H. Shen, S. A. Marion, J. Tsui and K. Teschke, "Pesticide exposure and risk of Parkinson's disease - a population-based case-control study evaluating the potential for recall bias," *Scand J Work Environ Health,* vol. 37, no. 5, pp. 427-436, 2011.

[34] M. Caballero, S. Amiri, J. T. Denney, P. Monsivais, P. Hystad and O. Amram, "Estimated exposure to agricultural chemicals and premature mortality by Parkinson's disease in Washington state," *Int J Environ Res Public Health,* vol. 15, p. 2885, 2018.

[35] C. B. Breckenridge, C. Berry, E. T. Chang, R. L. Sielken Jr and J. S. Mandel, "Association between Parkinson's disease, cigarette smoking, rural living, well-water consumption, farming, and pesticide use: systematic review and meta-analysis," *PLoS ONE,* vol. 11, no. 4, p. e0151841, 2016.

[36] J. K. Howard, "A clinical survey of paraquat formulation workers," *British Journal of Industrial Medicine,* vol. 36, pp. 220-223, 1979.

[37] J. K. Howard, N. N. Sabapathy and P. A. Whitehead, "A study of the health of Malaysian plantation workers with particular reference to paraquat spraymen," *British Journal of Industrial Medicine,* vol. 38, pp. 110-116, 1981.

[38] J. Brent and T. H. Schaeffer, "Systematic review of Parkinsonian syndromes in short- and long-term survivors of paraquat poisoning," *JOEM,* vol. 53, no. 11, pp. 1332-1336, 2011.

[39] J. R. Richardson, Y. Quan, T. B. Sherer, J. T. Greenamyre and G. W. Miller, "Paraquat neurotoxicity is distinct from that of MPTP and rotenone," *Toxicol Sci,* pp. 193-204, 2005.

[40] J. A. Tomenson and C. Campbell, "Mortality from Parkinson's disease and other causes among a workforce manufacturing paraquat: a retrospective cohort study," *BMJ Open,* vol. 2, p. e000283, 2011.

[41] N. J. Phillips, J. Reay and C. N. Martyn, "Validity of mortality data for Parkinson's disease," *J Epidemiol Community Health,* vol. 53, pp. 587-588, 1999.

[42] Brouwer, D H; et al, "Formulation R BIX (A9409AL): Measurement of excreted paraquat in workers following a single day's habitual use in Luzerne," SYNG-PQ-00022018, (TNO Study Code 6420-03), June 2007.

[43] Brouwer, D H; et al, "Meaasurement of excreted paraquat in workers following repeated day's habitual use in vines," SYNG-PQ-00125329 , (TNO Study Plan 6420-02), May 2007.

[44] Brouwer, D H; et al, "Formulation R BIX (A9409AL): Measurement of excreted paraquat in workers following three consecutive days habitual use in bananas," (TNO Study Plan P6420-5), SYNG-PQ-00124055, August 2007.

[45] Brouwer, D H; et al , "Formulation R BIX (A9409AL): Measurement of excreted paraquat in workers following a single day's habitual use in vineyards," (TNO Study Plan 6420-01), SYNG-PQ-00125211, May 2007.

[46] Brouwer, D H; et al , "Formulation R BIX (A9409AL): Measurement of excreted paraquat in workers following three consecutive days use in bananas with PPE according to label recommendations," (TNO Study Plan P6420-06), SYNG-PQ-00124616, September 2007.

[47] G. Chester and B. H. Woollen, "A study of the occupational exposure of Malaysian plantation workers to paraquat," Report No. CTL/P/580, SYNG-PQ-00693891, 12/19/1980.

[48] D. J. Meier, "Paraquat: worker exposure during mixing, loading, and application of Gramoxone Extra to pecans using vehicle-mounted, ground boom equipment," Report Number RR95-019B, SYNG-PQ-01806986, 4/28/1995.

[49] A. Swan, "Exposure of spray operators to paraquat," *Brit J Indust Med,* vol. 26, pp. 322-329, 1969; SYNG-PQ-00228608.

[50] J. K. Andersen and D. A. Di Monte, "Paraquat and Parkinson's disease (PD)," *Encyclopedia of the Neurological Sciences Vol. 3,* pp. 792-794, 2014.

[51] K. Ossowska, J. Wardas, M. Smialowska, K. Kuter, T. Lenda, J. M. Wieronska, B. Zieba, P. Nowak, J. Dabrowska, A. Bortel, A. Kwiecinski and S. Wolfarth, "A slowly developing dysfunction of dopaminergic nigrostriatal neurons induced by long-term paraquat administration in rats: an animal model of preclinical stages of Parkinson's disease?," *European Journal of Neuroscience,* vol. 22, pp. 1294-1304, 2005.