# ANALYSIS AND CRITIQUE OF THE CDC'S HANDLING OF THE THIMEROSAL EXPOSURE ASSESSMENT BASED ON VACCINE SAFETY DATALINK (VSD) INFORMATION

**Safe Minds**
**October 2003**

# SUMMARY

**The CDC's approach to analysis of the VSD database demonstrates a pervasive pattern of bias and conscious manipulation of samples, statistics and findings to produce a negative finding regarding the dangers of thimerosal exposure to children**

**Despite significant problems with study design and data quality and contrary to public statements by the CDC, the VSD analyses of autism, NDDs and speech delay provide support for a causal relationship between thimerosal exposure and childhood developmental disorders**

**Comparisons at a population level across HMOs suggest that compliance with the recommended vaccine schedule of thimerosal exposure was associated with high rates of neurological disorders and developmental delay.**

**Full compliance populations reported to HMOs disease frequencies exceeding 5% of the birth populations. Extrapolating these rates to a national level suggests that the population harmed by thimerosal exposure may number in the millions**

# SUMMARY
# Issue of Study Design

**Starting in late 1999, the CDC developed and then modified its VSD study protocol many times. By November 2003, the report had gone through four generations of modification**

**At each generation, the research team made subtle but powerful changes to its original study protocol. With each change, troubling findings were obscured or made less significant.**

- **in the early generations, the results were so troubling that the principal investigator stated privately his opinion that thimerosal exposure had caused harm**
- **by the later generations, the authors had concluded that any troubling findings could be dismissed as the result of random chance**

**Sufficient flexibility was available to the VSD research team to make such modifications. The design parameters of the thimerosal study were numerous and highly technical. But the general drift of their design changes was clear, to reduce statistical power through conscious manipulation of statistical methods, data classification and samples.**

# SUMMARY
# Tracing CDC's Findings Across Study Generations

**Despite these attempts at distortion, several conclusions can be reached from the data with respect to the risk of thimerosal exposure from vaccines**

- **autism outcomes for the highest exposure category (at 3 months and relative to lowest exposure group) showed high relative risks viewed in context of the limits of the data sources and study methods**
  - **2.48 in first generation of the analysis, using Cox model, and meets legal standard for causality**
  - **2.15 in second generation using actual data, and statistically significant**
- **Neurodevelopmental disorders showed increased and statistically significant odds ratios and dose response curves for sample populations with sufficient statistical power at all times in which results were reported**
  - **dose response relationship consistently showed 99% confidence**
  - **reported relationships were stable across smaller samples as well**
- **Significant findings for developmental speech and language delay persisted across study generations, especially in the largest sample groups**

# SUMMARY
# Thimerosal Effect at a Population Level

**The magnitude of the potential harm from thimerosal can be demonstrated by a simple analysis, comparing exposures and outcomes across HMOs at the population level**

- **the HMOs with higher vaccine schedule compliance and with higher thimerosal exposure report higher rates of childhood developmental disorders**
- **these elevated rates reach high absolute levels, with rates over 5% for the highest exposure populations**

**These elevated rates are likely understated, since the study populations were generally young and under-diagnosed and many of these disorders may not have been managed by or reported to medical care providers**

**Assuming that high levels of thimerosal exposure prevailed in the entire decade of the 1990s. then roughly 40 million children born during the decade were at risk of harm from thimerosal exposure. If rates of harm exceeded 5%, then over 2 million children may have been measurably harmed by the mercury exposures to which thimerosal-containing vaccines contributed**

# CONTENTS

# ISSUES OF VSD THIMEROSAL STUDY DESIGN

**Despite looming controversy, the original study protocol took pains to point out the inherent noise in the VSD sample and explained that this created "biases toward the null of the relative risk." To compensate, the protocol**

- **envisioned sample pooling and diagnostic grouping to enhance statistical power**
- **provided for the active review of moderate risk outcomes**
- **identified numerous limitations of the automated databases**

**The specific work plan required research design decisions on numerous dimensions, each of which offered potential to affect study outcomes**

- **dimensions of the statistical model: number and choice of strata**
- **specific measures for reporting: odds ratios vs. dose-response curves**
- **population parameters: inclusion criteria, age distribution, follow up**
- **pooling vs. separation of HMO sample populations**
- **grouping vs. separation of diagnostic outcomes**
- **classification standards for diagnosis and diagnostic categories**

**Across multiple iterations of design modifications, the research team produced four generations of reports. These modifications consistently pointed the research in ways that weakened the sensitivity of the statistical analysis**

# STUDY PROTOCOL WAS INITIATED AGAINST BACKGROUND OF LATENT OPPOSITION IN THE PUBLIC HEALTH COMMUNITY

**"I am really concerned that we have taken off like a boat going down one arm of the mangrove swamp at high speed, when in fact there was not enough discussion really early on about which way the boat should go at all. And I really want to risk offending everyone in the room by saying that perhaps this study should not have been done at all, because the outcome of it could have, to some extent, been predicted, and we have all reached this point now where we are left hanging, even though I hear the majority of consultants say to the Board that they are not convinced there is a causality direct link between thimerosal and various neurological outcomes. I know how we handle it from here is extremely problematic.**

- **Dr. John Clements, WHO representative to Simpsonwood discussions, June 7, 2000**

# BACKGROUND: THE VACCINE SAFETY DATALINK AND THE THIMEROSAL SAFETY ANALYSIS

## Vaccine Safety Datalink

**A collaborative project between NIP, CDC and HMO's established in 1990.**

**Database includes approximately 6 million clients representing 2% of the U.S. population.**

**Database incorporates information on vaccination, medical outcome and co-variant data.**

**It is used to monitor vaccine safety issues**

**Only a single independent research group has ever been granted (highly constrained) access**

## Thimerosal Safety Analysis

**In 1999 the FDA acknowledged that infants may have been exposed to mercury, a known neurotoxin, in excess of Federal safety guidelines.**

**This exposure was a result of a preservative, Thimerosal, which contained ethyl mercury, utilized in some, but not all infant vaccines.**

**In light of these concerns, the CDC began to investigate the impact of vaccine thimerosal exposure utilizing the Vaccine Safety Database**

**Only two of the participating HMO's had outpatient records necessary for the investigation**

- **Group Health Cooperative (GHC)**
- **North California Kaiser (NCK).**

# SELECTED ELEMENTS OF THE ORIGINAL STUDY PROTOCOL

Case definition:

Case conditions will be determined from the phase I results and will include any outcome that showed a potential relationship to the exposure. A relationship will be considered plausible if statistically significant or a relative risk of 1.5 or higher is found. This would allow weak suggestive findings to be further investigated as we expect a bias towards the null of the relative risk, caused by the lack of sensitivity of the automated data.

Statistical analysis:

Crude odds ratios will be calculated using 2-way contingency tables.

Adjusted odds ratios will be derived from conditional logistic regression models whereby the estimated OR and it's standard error will be corrected for the two-stage sampling method.

Sample size and power:

To be determined for specific conditions and the available number of cases. A power of 80% to show a relative risk of at least 2.0 with a precision of 95% will be aimed for. The total number of cases encountered in children of all ages in the VSD automated data is included for each condition in the preliminary list for phase I.

# ACKNOWLEDGED LIMITATIONS OF THE STUDY DATABASES

**Possible misclassification of exposures**

- **hepatitis B birth dose may not have been recorded (up to 40% of sample, based on recorded 1 month exposure levels)**
- **thimerosal free Hib vaccines were recorded as thimerosal-containing**

**Misclassification of outcomes in an automated database: ICD9 codes**

- **large number of cases (158) of an unusual disorder: "misery disorder" reported in early generations of the report, exceeding reported autism cases**
- **many cases of autism might be initially diagnosed as another neuro-developmental disorder, such as speech delay, language delay or misery (?)**
- **low ascertainment rates for autism among younger children due to typical delays in obtaining autism diagnoses (median age at diagnosis is 4-5 years, yet sample included children starting at 1 year of age)**

**Medical care utilization factors, for which little evidence was apparent until the final study**

**Only conditions that come to medical attention, creating a bias to understate the true incidence of disorders**

**Insufficient power for some conditions and numerous risk calculations**

- **zero exposure populations were small, cases at those levels were often proportionately smaller and therefore aggregated with low exposures**

# SELECTED ADDITIONAL LIMITATIONS OF STUDY DATABASES

**Serious problems with Harvard Pilgrim, the HMO used for Phase 2 testing**

- **small population size, less than 15% of largest HMO**
- **diagnostic classifications not based on International Classification of Diseases (ICD) standards but instead based on "Costar" codes, an entirely different classification scheme**
- **extremely low exposure variation, combined with small sample sizes, provided virtually no data on low exposure populations**
- **severe financial distress (including bankruptcy) during the study period in 1999-2000 accompanied by publicized concerns over information integrity, payments and record quality**

**Large discrepancies between recorded cases and "chart-confirmed" cases**

- **leading to large exclusions in later generations following extensive chart review**
- **exclusion decisions, however, were based on the credentials of the diagnostician rather than the quality of the diagnosis itself**

# DESIGN CHOICES WERE MODIFIED TO INFLUENCE OUTCOMES OF LATER REPORT GENERATIONS

| | Early design intent | Later design choices |
|---|---|---|
| **Strata for "conditional logistic regression analysis"** | ***From over 100 strata***<br>**HMO site**<br>**Year *and* month of birth**<br>**"Adjusted for gender"** | ***To even greater stratification***<br>**Separated HMOs completely**<br>**Added clinics within largest HMO**<br>**Year, month and gender retained** |
| **Specific measures** | ***From transparent measures***<br>**"Crude risk ratios" (not used)**<br>**"Adjusted" odds ratios (OR)**<br>**• at each exposure level**<br>**Dose-response curves** | ***To low transparency analyses***<br>**Dose-response curves**<br>**OR for grouped exposures only** |
| **Population exclusions** | ***From original exclusion standard***<br>**Age and congenital exclusions** | ***To highly variable exclusions***<br>**Across reports and HMOs** |
| **Diagnostic groupings** | ***From large set and groups***<br>**15 diagnoses in 3 groups**<br>**Broad category most significant** | ***To small set and no groups***<br>**11 or fewer diagnoses**<br>**Broad category not reported** |
| **HMO data pooling** | ***From pooled HMO data with strata for site adjustments*** | ***To complete separation of all HMO results, with reduced power*** |

# DESIGN ITERATIONS PLAYED OUT THROUGH FOUR GENERATIONS OF FORMAL REPORTS

***Generation 1*: first internal draft, obtained through FOIA request**

- **February 29, 2000**
- **Phase 1 only: 2 HMOs reported jointly**
  - **Group Health Cooperative**
  - **Northern California Kaiser**

***Generation 2*: first disseminated draft, discussed at ACIP and Simpsonwood**

- **June 2000**
- **Includes added Phase 2 HMO**
  - **Harvard Pilgrim, reported separately**

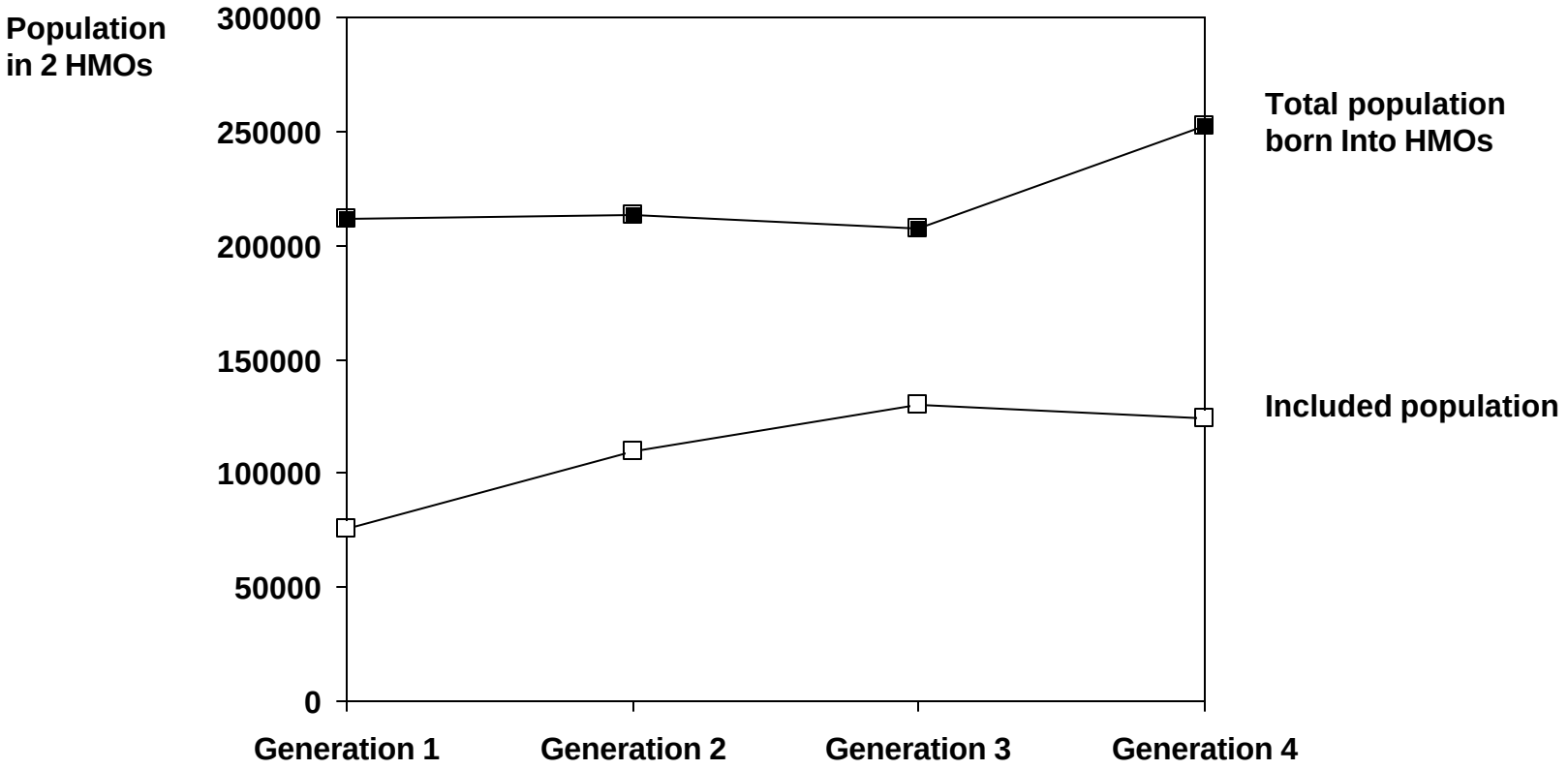
***Generation 3*: first public draft, presented at IOM hearing**

- **July 2001**
- **3 HMOs reported separately**

***Generation 4*: first published report in Pediatrics**

- **November 2003 (submitted January 2003, accepted July 2003)**
- **3 HMOs reporting separately, with additional stratification at largest HMO**

# STUDY POPULATIONS CHANGED IN EACH GENERATION WITH TOTAL AND INCLUDED POPULATIONS OFTEN MOVING IN OPPOSITE DIRECTIONS



# DIAGNOSTIC INCLUSIONS HAVE CHANGED IN SOME KEY CATEGORIES OVER REPORT GENERATIONS

| | | Generation 1 | Generation 2 | Generation 3 | Generation 4 |
|---|---|---|---|---|---|
| 299.0 | autism | X | X | X | X |
| 299.8 | other childhood psychoses | X | X | X | X |
| 299.9 | other unspecified psychosis | X | - | - | - |
| 307.0 | stammering and stuttering | X | X | X | X |
| 307.2 | tics | X | X | X | X |
| 307.4 | sleep disorders | X | X | X | X |
| 307.5 | eating disorders | X | X | X | X |
| 313.1 | misery disorder | X | X | - | - |
| 313.5 | emotional disturbances | - | - | X | X |
| 313.8 | mixed emotional | X | X | - | - |
| 314.0 | attention deficit syndrome | X | X | X | X |
| 315.0 | specific delays in development | X | - | - | - |
| 315.31 | developmental language delay | X | X | X | X |
| 315.39 | developmental speech delay | X | X | X | X |
| 315.3 | speech or language delay | - | - | - | X |
| 315.4 | coordination disorder | - | - | X | X |
| 315.9 | unspecified developmental delay | X | X | - | - |
| | all neurological disorders (NDDs) | X | X | X | - |

# BIOLOGICAL PLAUSIBILITY REINFORCED EARLY CONCERNS AND UNDERSCORED SENSITIVITY OF ASSESSMENTS TO SMALL DIFFERENCES IN EXPOSURES AND OUTCOMES

**" When I saw this, and I went back through the literature, I was actually stunned by what I saw because I thought it is plausible. First of all there is the Faeroe study, which I think people have dismissed too easily, and there is a new article in the same Journal that was presented here, the Journal of Pediatrics, where they have looked at PCB. They have looked at other contaminants in seafood and they have adjusted for that, and still mercury comes out. That is one point. Another point is that in many of the studies with animals, it turned out that there is quite a different result depending on the dose of mercury. Depending on the route of exposure and depending on the age at which the animals were exposed. Now, I don't know how much you can extrapolate that from animals to humans, but that tells me mercury at one month of age is not the same as mercury at three months, at 12 months, prenatal mercury, later mercury. There is a whole range of plausible outcomes from mercury. "**

- **Dr. Verstraeten, commenting in Simpsonwood discussions, June 7, 2000**

# CONTENTS

**Issues of VSD thimerosal study design**

**Tracing the CDC's findings across report generations**

- **autism**
- **neurological developmental disorders (NDDs)**
- **speech/language delay**

**Assessment of thimerosal effect at a population level across HMOs**

# AN INSIDER'S VIEW OF THE VSD FINDINGS

**"I put four [on a scale of 1-6 of probable causality] and I did so for a number of reasons.**

**"The number of dose-related relationships are linear and statistically significant. You can play with this all you want. They are linear. They are statistically significant.**

**"The positive relationships are those that one might expect from the Faeroe Islands studies. They are also related to those data we do have on experimental animal data and similar to the neurodevelopmental tox data on other substances, so I think you can't accept that this is out of the ordinary. It isn't out of the ordinary.**

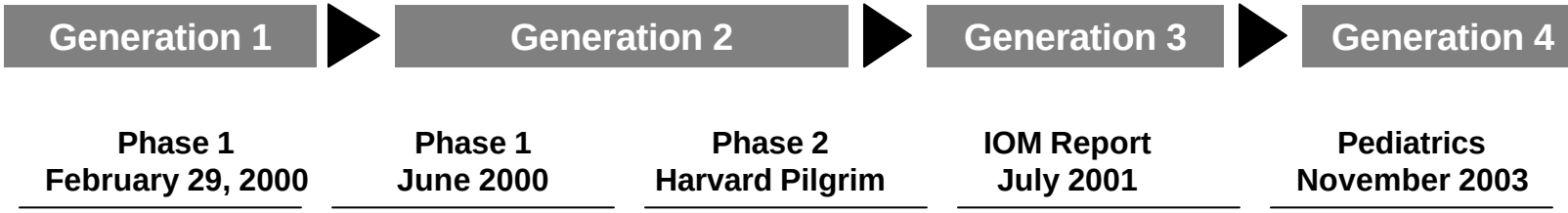
**"The Seychelles Islands studies, and somebody said the Faeroe Islands studies both, were chronic exposures. We are not talking necessarily about chronic exposure. We are talking about a series of acute exposures and at one point in time that exposure is much greater on one day than any of the Seychelles Islands.**

**"The increased incidence of neurobehavioral problems in children in the past few decades is probably real…I work in the school system where my effort is entirely in special education and I have to say that the number of kids getting help in special education is growing nationally and state by state at a rate we have not seen before. So there is some kind of increase…**

**"The rise in frequency of neuro-developmental disorders whether it is ascertainment or real is… much too graphic. We don't see that kind of genetic change in 30 years."**

- **Dr. William Weil, pediatrician and expert panelist, commenting in Simpsonwood discussions, July 7, 2000**

# TRACING THE CDC'S APPROACH TO THIMEROSAL RISK ANALYSIS REQUIRES COMPARISON OF FINDINGS ACROSS GENERATIONS



# 1. TRACING THE AUTISM FINDINGS ACROSS REPORT GENERATIONS

**Generation 1 analysis found an elevated risk of autism in highest exposure group**

- **2.48 relative risk in >62.5 mcg exposure group ("on schedule" children) as compared to <37.5 mcg exposure group**
- **finding not quite statistically significant due to wide 95% confidence intervals**
- **risk level meets a legal standard of proof, exceeding 2.0**

**Elevated risk finding resulted largely from two of five of sample years (four and five year olds), since youngest children (3 years old and younger) show sharply reduced ascertainment levels, thereby**

- **increasing statistical noise in the autism sample: many "false negative" autism cases among younger group**
- **reducing the strength of the signal generated by the higher ascertainment group**

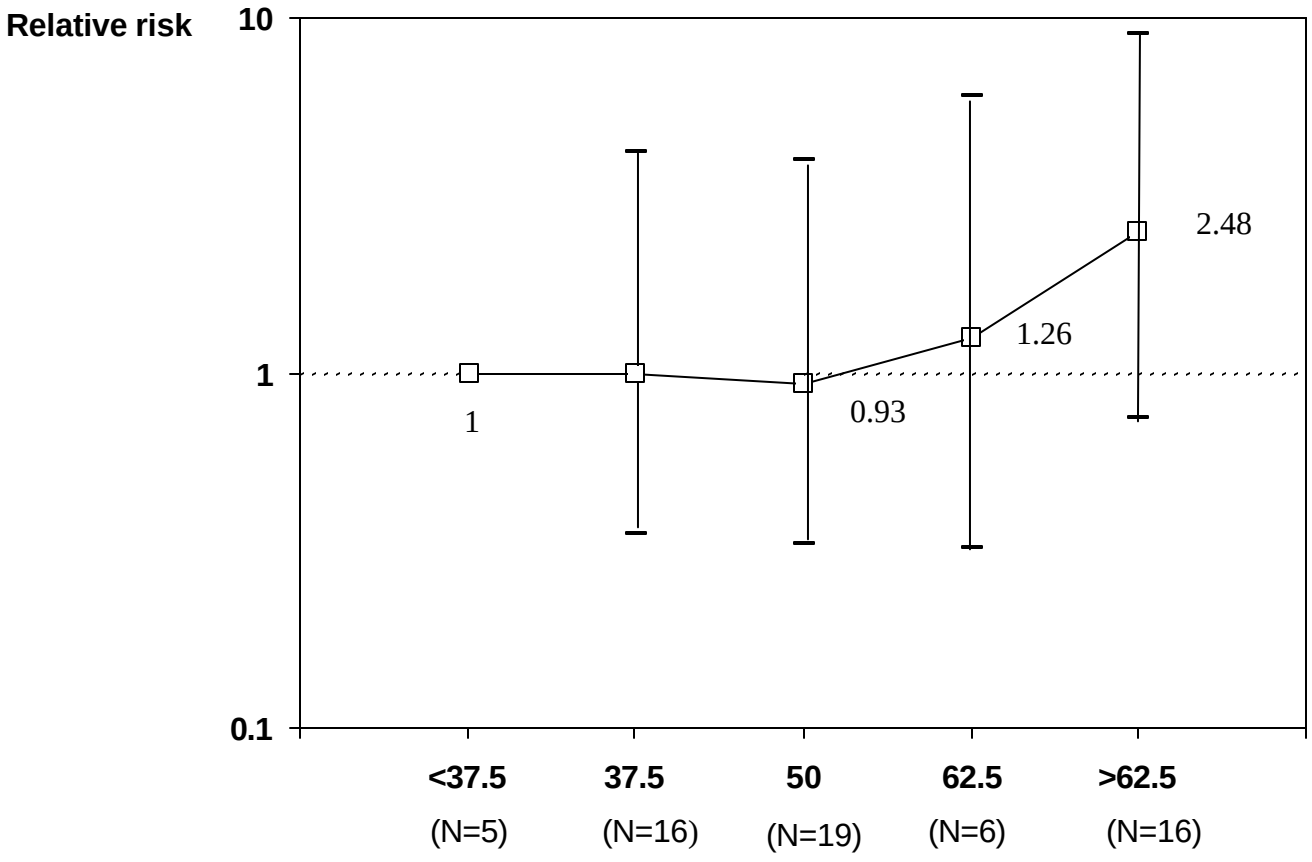
**The sample for the Generation 2 analysis augmented the Generation 1 sample by reducing exclusions in two areas**

- **"continuously enrolled first year", adding more younger children**
- **"no congenital, birth disorder", adding children initially excluded due to congenital disorders**

**The new groups reduced the modeled autism risk below 2.0, from 2.48 to 1.69 in Generation 2**

- **without the Cox model, the odds ratio remained at 2.15 and was statistically significant**

# RELATIVE RISK OF AUTISM FROM MERCURY EXPOSURE AT THREE MONTHS OF AGE: GENERATION 1 FINDINGS



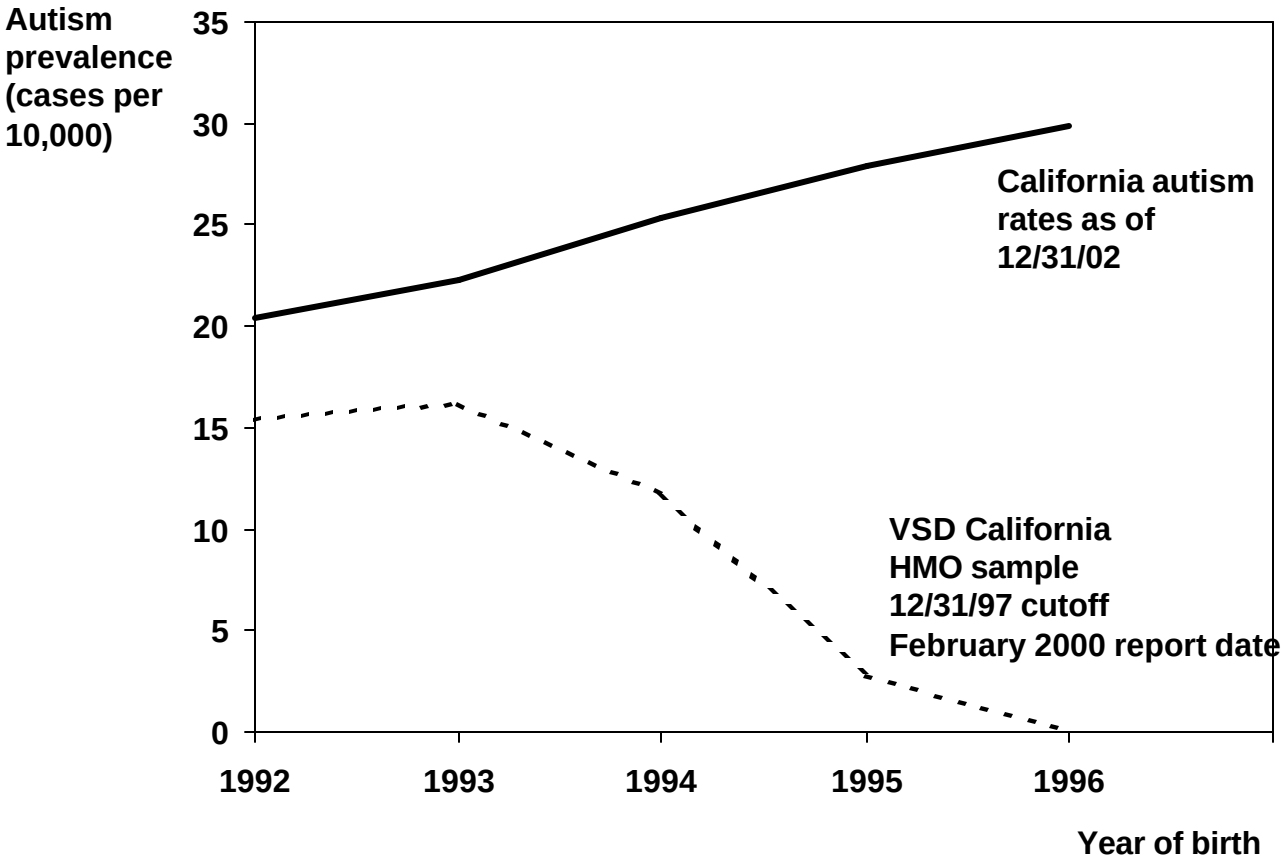
# GENERATION 1 AUTISM ANALYSIS RELIED ON A SAMPLE WITH SHARPLY REDUCED ASCERTAINMENT IN 1-3 YEAR OLDS

The endpoint was defined as the first of the following dates:

- the date of first diagnosis
- the first date that a child stopped being enrolled in the HMO
- December 31st, 97

| | | | T | Year of birth (%) | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | 92 | 93 | 94 | 95 | 96 |
| ALL kids | | 116229 | 76509 | 17 | 21 | 19 | 19 | 22 |
| Neurologic degenerative disorders: | | 145 | 57 | 21 | 19 | 33 | 7 | 12 |
| 330.x | Cerebral degenerations usually | 7 | 4 | 25 | 25 | 0 | 0 | 50 |
| 331.x | Other cerebral degenerative disease | 59 | 14 | 14 | 29 | 29 | 7 | 21 |
| 333.x | Other extrapyramidal disease and | 47 | 28 | 29 | 18 | 29 | 4 | 11 |
| 334.x | Spinocerebellar disease | 8 | 3 | 33 | 0 | 33 | 33 | 0 |
| 335.x | Anterior horn cell disease | 3 | 3 | 33 | 33 | 33 | 0 | 0 |
| Neurologic developmental disabilities: | | 2991 | 1743 | 22 | 27 | 24 | 17 | 8 |
| 299.0 | Autism | 109 | 67 | 30 | 39 | 25 | 6 | 0 |
| 299.x | Other childhood psychosis | | | | | | | |

# ASCERTAINMENT BIAS IN GENERATION 1 AUTISM SAMPLE AS COMPARED TO KNOWN CALIFORNIA PREVALENCE RATES




Sources: VSD analysis, 2/29/00, California DDS

# AUTISM POPULATION CHANGES FROM GENERATION 1 TO GENERATION 2

| | 2/29/00 | June 2000 | Difference (cumulative)[1] |
|---|---|---|---|
| "Born into GHC or NCK" 1992-97 | 211,693 | 213,185 | 1,492 |
| "Continuously enrolled first year“ | 121,441 | 142,264 | 20,832 |
| 1 polio shot by year 1 | 116,687 | 139,994 | 22,447 |
| Not premie | 111,239 | 132,391 | 21,152 |
| No maternal hep B Ig | 111.047 | 132,114 | 21,067 |
| No congenital /birth disorder | 75,659 | 109,993 | 34,334 |

(1) inclusion differences are cumulative in descending fashion down the third column. Total inclusion differences are 34,334

# “CRUDE” CALCULATIONS DEMONSTRATE SIGNIFICANT AUTISM RISK USING GENERATION 2 DATA

| | |
|---|---|
| **Low* exposures: <37.5 micrograms (3 months)** | |
| • cases | 11 |
| • population | 12,429 |
| • rate per 10,000 | 8.85 |
| **High exposures: >75 micrograms (3 months)** | |
| • cases | 28 |
| • population | 14,739 |
| • rate per 10,000 | 19.0 |
| **Odds ratio** | |
| • crude | 2.15 |
| • Cox model adjusted | 1.69 |

***Note: Low exposures defined as <37.5 mcg due to absence of reported data on zero exposure level**

# DISMISSING AUTISM RISK

| | Generation 1 | Generation 2 | | Generation 3 | Generation 4 |
|---|---|---|---|---|---|
| | Phase 1<br>February 29, 2000 | Phase 1<br>June 2000 | Phase 2<br>Harvard Pilgrim | IOM Report<br>July 2001 | Pediatrics<br>November 2003 |
| **Autism risk** | **2.48 (n.s.)** | **1.69 (n.s.)**<br>**•2.15 "crude"** | **not done** | **n.s.** | **n.s.** |
| **Neurological developmental disorders risk (NDDs)** | | | | | |
| **Speech/language delay risk** | | | | | |

# 2. TRACING THE NDD FINDINGS ACROSS REPORT GENERATIONS

**Phase 1 analyses in both Generation1 and Generation 2 analyses found significant increased risk of neurological developmental delay due to mercury exposure**

- **significant elevated risks of 1.59 and 1.64**
- **significant trend (dose-response) with increased risk with rising exposure**

**Phase 2 analysis in Harvard Pilgrim was restricted to three specific sub-categories and did not attempt to test the finding in the umbrella NDD category**

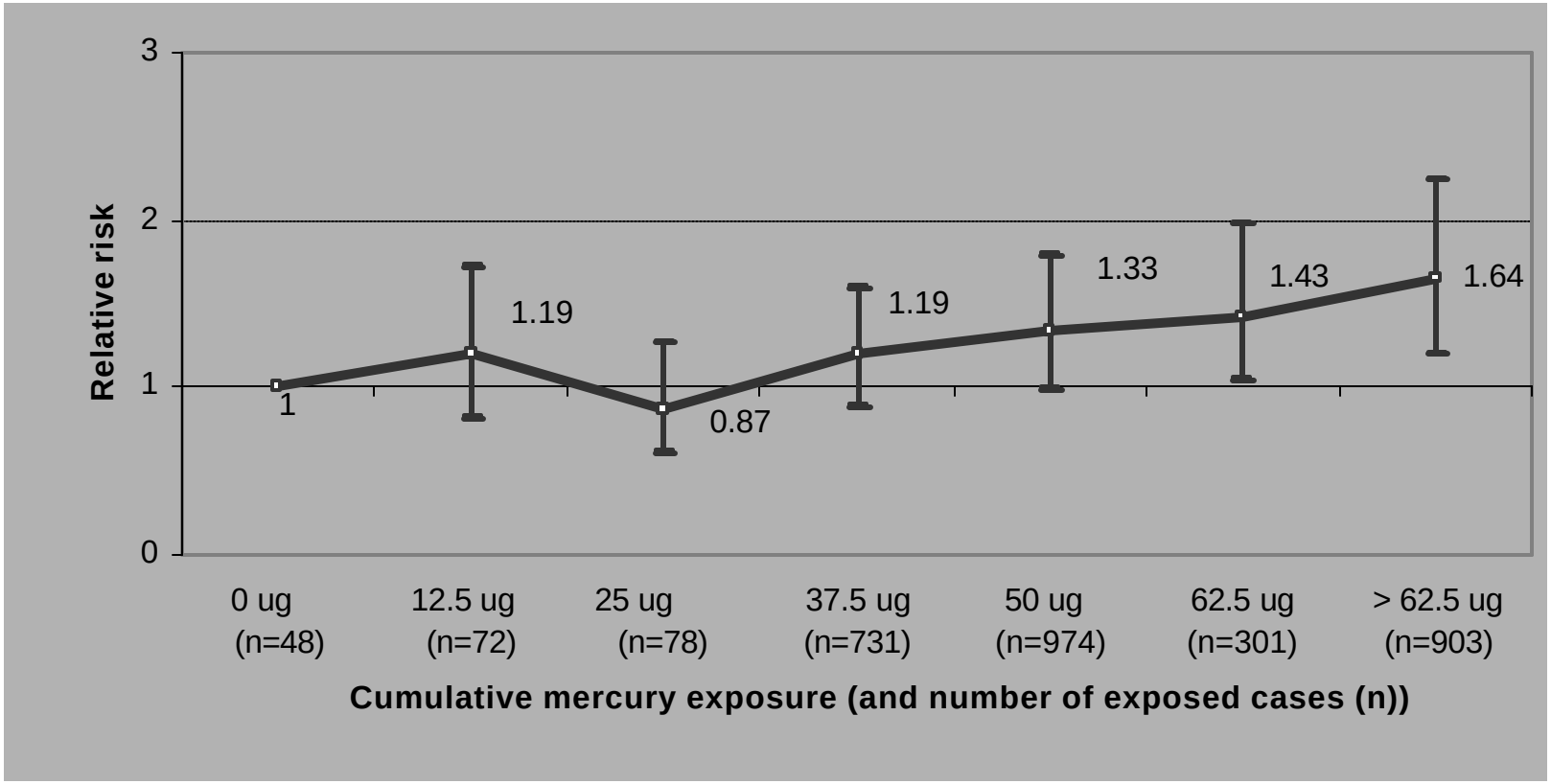
**Generation 3 analysis for IOM report included a new analysis, splitting the HMOs into two separate groups, thereby reducing statistical power of all analyses**

- **reducing sample size, both overall and in the smaller HMO**
- **finding no significant risk or trend in the smaller HMO, with 13% of the population of the larger HMO**

**Generation 3 findings reported that "results were not consistent across HMOs"**

**The Generation 4 report omitted discussion of NDDs entirely**

# RELATIVE RISK FROM MERCURY EXPOSURE AT 3 MONTHS OF AGE: NEUROLOGIC DEVELOPMENTAL DISORDERS (NDD): GENERATION 2 FINDINGS




Trend (RR per added microgram):
1.007 (1.004, 1.010), p < 0.01

# EFFECT OF SPLITTING THE HMOS ON THE NDD RESULT IN GENERATION 3

| | HMO A | HMO B |
|---|---|---|
| Actual HMO | Group Health Cooperative | Northern California Kaiser |
| Population size | 15,309 | 114,966 |
| Full compliance rate at three months with CDC vaccination schedule (% at 75 mcg) | 60% | 15% |
| NDD finding | n.s. | Statistically significant trend ($p<0.01$) |

# “CRUDE” CALCULATIONS DEMONSTRATE FAR HIGHER NDD RISK USING GENERATION 2 DATA

| | |
|---|---|
| **Low exposures, 0 micrograms (3 months)** | |
| • cases | 48 |
| • population | 4,510 |
| • rate per 10,000 | 106.4 |
| **High exposures >75 micrograms (3 months)** | |
| • cases | 903 |
| • population | 14,739 |
| • rate per 10,000 | 612.7 |
| **Odds ratio** | |
| • crude | 6.1 |
| • Cox model adjusted | 1.64 |

# DISMISSING NDD RISK

| | Generation 1 | Generation 2 | | Generation 3 | Generation 4 |
|---|---|---|---|---|---|
| | Phase 1<br>February 29, 2000 | Phase 1<br>June 2000 | Phase 2<br>Harvard Pilgrim | IOM Report<br>July 2001 | Pediatrics<br>November 2003 |
| **Autism risk** | 2.48 (n.s) | 1.69 (n.s.)<br>•2.15 "crude" | not done | n.s. | n.s. |
| **Neurological developmental disorders risk (NDDs)** | 1.59 | 1.64<br>•6.1 "crude"<br><br>1.007 per mcg trend<br>(p<0.01) | not done | "results not consistent between HMOs A and B" | not reported |
| **Speech/language delay risk** | | | | | |

# 3. TRACING THE SPEECH/LANGUAGE DELAY FINDINGS ACROSS REPORT GENERATIONS

**The speech delay analysis in Generation 2 found a significant dose response relationship between mercury exposure and developmental speech and language delays in the pooled phase 1 HMOs**

**Phase 2 analysis in Generation 2 report included Harvard Pilgrim, an HMO with a significantly smaller population and tested only speech delay**

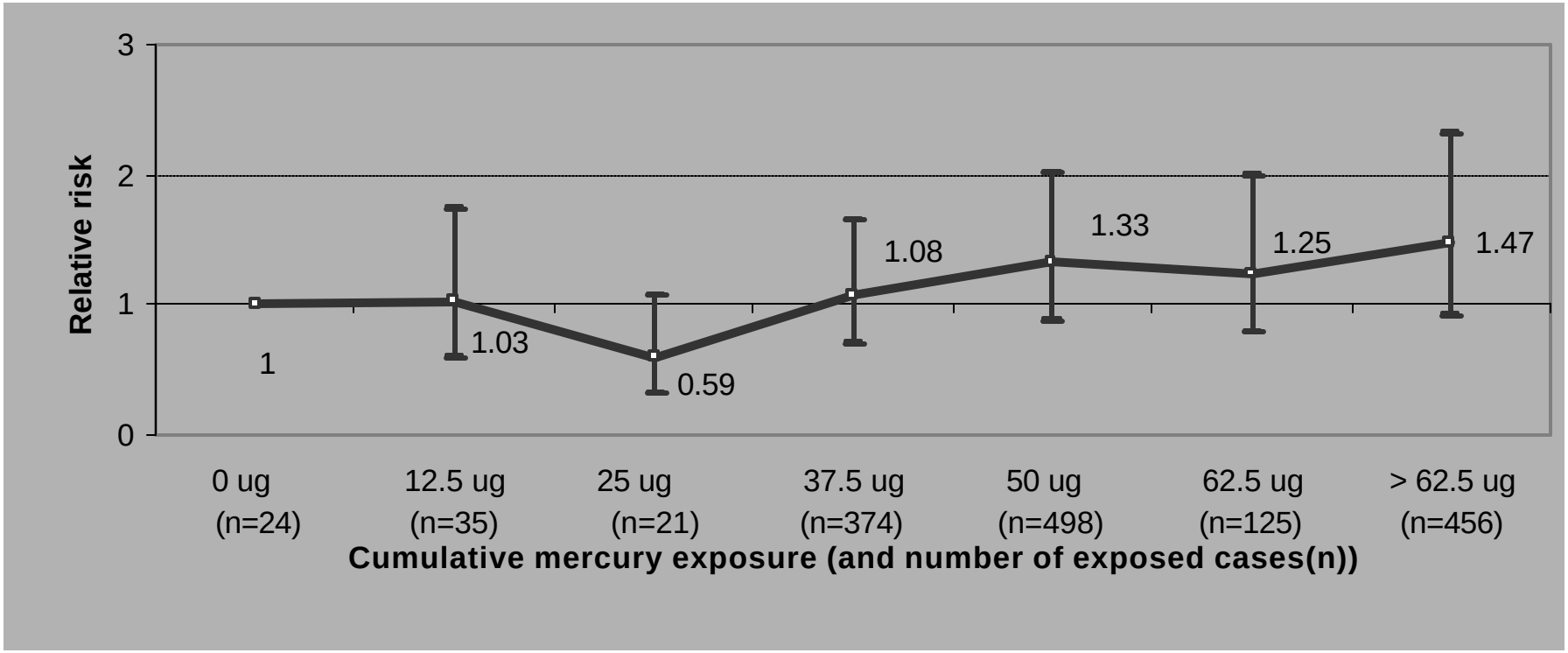
- **15% the size of Northern California Kaiser**
- **different diagnostic classification scheme: Costar vs. ICD 9**

**Generation 3 report dismissed the speech and language delay findings by unpooling the two HMOs**

- **“results not consistent between phases”**
- **“results not consistent between HMOs”**
- **but these dismissals were based on reduced sample sizes and analyses with reduced statistical power**

**Generation 4 report reduced significance of remaining NCK findings by stratifying based on clinic, claiming that “chance alone” has produced significant findings**

# RELATIVE RISK OF SPEECH DELAY FROM MERCURY EXPOSURE AT THREE MONTHS OF AGE
## Generation 2 Findings




Trend (RR per added microgram):
1.008 (1.004, 1.013), p = 0.0004

# EFFECT OF SPLITTING THE HMOs ON SPEECH DELAY RESULT
## Generation 3 Findings

| | HMO A | HMO B | HMO C |
|---|---|---|---|
| Actual HMO | Group Health Cooperative | Northern California Kaiser | Harvard Pilgrim |
| Population size | 15,309 | 114,966 | 17,547 |
| Full compliance rate at three months | 60% | 15% | 65% |
| Language delay | n.s. | 1.20 per 12.5 mcg trend at 3 months ($p<0.01$) | not done |
| Speech delay | n.s | 1.10 per 12.5 mcg trend at 1 month ($p<0.05$) | n.s. |

# NEW STRATIFICATION BY CLINIC SITE CREATES "AN APPRECIABLE CONFOUNDER" FOR HMO B AND DEGRADES SIGNIFICANCE
## Generation 4 Findings

| | Generation 3 | Generation 4 | Comments |
|---|---|---|---|
| **Developmental speech delay** | **(risk per 12.5 mcg)** | **(risk per 12.5 mcg)** | **"Our study encompassed a large number of separate analyses and, by chance alone, at least some associations would be expected to be statistically significant"** |
| **•0-1 months** | **1.10*** | 1.02 | |
| **•2-3 months** | 1.06 | 1.04 | |
| **•4-5 months** | 1.00 | - | |
| **•6-7 months** | **1.06*** | 1.02 | |
| **•0-7 months** | **1.05**** | **-** | |
| **Developmental language delay** | | | **"HMO B is the only HMO in our study where speech therapy is not covered by the health plan"** |
| **•0-1 months** | **1.37**** | 1.06 | |
| **•2-3 months** | **1.20**** | **1.13*** | |
| **•4-5 months** | 1.02 | **-** | |
| **•6-7 months** | 1.07 | **1.07*** | |
| **•0-7 months** | **1.09**** | **-** | |
| **Language or speech delay** | | | |
| **•0-1 months** | - | 1.03 | ***: p<0.05** |
| **•2-3 months** | - | 1.05 | ****: p<0.01** |
| **•6-7 months** | - | 1.02 | |

# DISMISSING SPEECH/LANGUAGE DELAY RISK

| | Generation 1 | Generation 2 | | Generation 3 | Generation 4 |
|---|---|---|---|---|---|
| | Phase 1<br>February 29, 2000 | Phase 1<br>June 2000 | Phase 2<br>Harvard Pilgrim | IOM Report<br>July 2001 | Pediatrics<br>November 2003 |
| Autism risk | 2.48 (n.s) | 1.69 (n.s.)<br>•2.15 "crude" | not done | n.s. | n.s. |
| Neurological developmental disorders risk (NDDs) | 1.59 | 1.64<br>•6.1 "crude"<br><br>1.007 per mcg dose-response (p<0.01) | not done | "results not consistent between HMOs A and B" | not reported |
| Speech/language delay risk | n.s. | 1.008 per mcg dose-response (p=0.0004) | n.s. | "results not consistent between phases" | "chance alone" would yield som positive associations |

# ONE AUTHOR OPENLY DISCUSSED POSSIBLE MANIPULATIONS

***Dr. Philip Rhodes (National Immunization Program, CDC) was assigned to look for ways to distort the methodology***

**"I also wanted to try to take a different look at the data because I think sometimes we make choices in our analyses. We conceptualize the problem very quickly and then everything else kind of depends on those initial choices and we don't always go down other pathways."**

***He made arguments to exclude the lowest exposure cases, claiming that the fact that their exposures were low suggested family behavior that made them unusual. The low rate of outcomes in this group, of course added significance to the statistical "signal" of a causal connection between thimerosal and harm***

**"I am not advocating totally throwing them [the low mercury exposure group] away and never considering them in any analysis, but at least for now let's think if we can establish if there are differences in this group of 37 to 75, then in a sense we really don't need them."**

***He made arguments to exclude some cases that had unusually high exposures and outcomes at the same time. Any high exposure, high outcome group would support the signal.***

**"There is an odd, small group of kids that supposedly receives separate DTP and Hib (note: with more thimerosal) and an unusually high percentage of those kids are outcomes…For example, if 1,500 kids were receiving one vaccine combination in that month of birth and 20 were receiving some other, I have removed the 20 completely from the analyses.**

***He made arguments to include non-comparable cases, all of which would serve to add "noise" that could obscure the signal.***

**"Now I take all those kids that Tom has excluded based on prematurity exclusion codes and throw them in. At one month I think there is some argument that is overdoing it. Throwing them all back in. I think there is a clear argument that is going too far, but that further brings things down."**

# THESE EARLY ATTEMPTS FAILED TO PRODUCE THE DESIRED RESULTS, BUT CAST IN DOUBT THE INTEGRITY OF THE ANALYTICAL PROCESS AND INSTITUTIONAL GOVERNANCE

***"So you can push, I can pull. But there has been substantial movement from this very highly significant result down to a fairly marginal result."***

- **Dr. Philip Rhodes, Pediatrics study co-author, CDC employee (National Immunization Program), speaking at Simpsonwood meeting, June 7, 2000**

# CONTENTS

**Issues of VSD thimerosal study design**

**Tracing the CDC's findings across report generations**

**Assessment of thimerosal effect at a population level across HMOs**

# CDC'S PRINCIPAL INVESTIGATOR COMMENTS ON CDC BIASES AND LIKELY CAUSALITY BEFORE PUBLIC DISCLOSURE

**"It is sort of interesting that when I first came to the CDC as an NIS officer a year ago only, I didn't know what I wanted to do, but one of the things I knew I didn't want to do was studies that had to do with toxicology or environmental health...Now it turns out that other people also thought that this study was not the right thing to do. So what I will present to you is the study that nobody thought we should do...**

**"Personally, I have three hypotheses. The first hypothesis is it is parental bias. The children that are more likely to be vaccinated are more likely to be picked up and diagnosed. Second hypothesis, I don't know. There is a bias I have not yet recognized, and nobody has yet told me about it.**

**"Third hypothesis. It's true. It's thimerosal."**

- ***Thomas Verstraeten, CDC analyst on VSD thimerosal study (now employed by GSK vaccine division), June 7, 2000***

# WHAT IF IT IS TRUE?
# DOSE-RESPONSE EFFECTS AT A POPULATION LEVEL

**Generation 3 report shifted to an approach of separate HMO analyses**

- **reducing sample sizes and statistical power**
- **relying on smallest samples with least exposure variation for negative findings**
- **assuming that differences in local diagnostic and health-care seeking practices across HMOs justified comparisons only within HMOs**
- **but failing to measure broad category of NDDs as a way to control for diagnostic practices**
- **also failing to measure dose response effects on a population basis across HMOs**

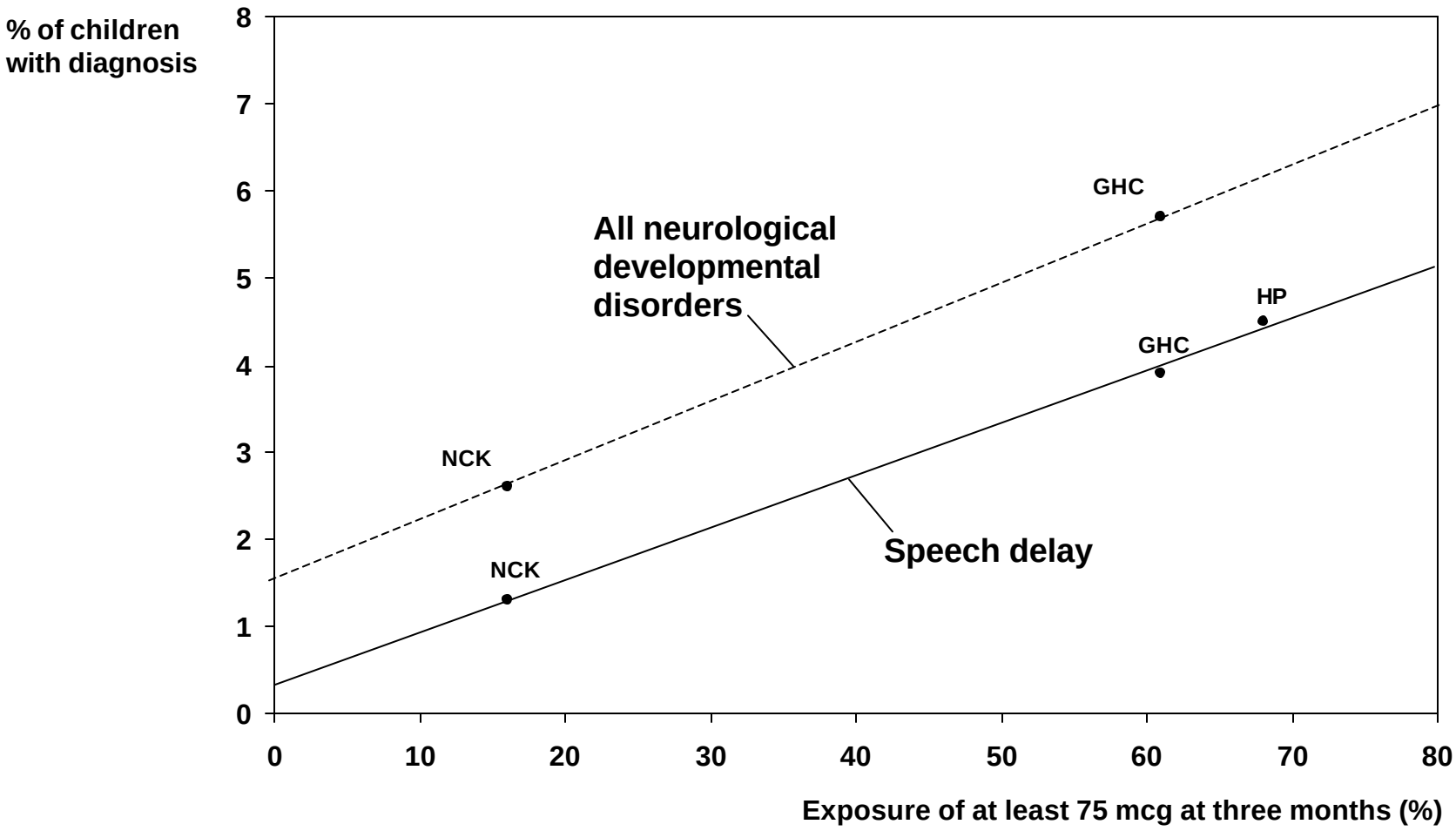
**At a population level, comparison across 3 HMOs demonstrates a dose response effect**

- **HMOs with highest vaccine compliance rates show highest frequency of NDD and speech delay**
- **no evidence that compliance differences result from health-care seeking behavior, with non-NDD diagnoses showing negative correlation with thimerosal exposure**

# WHAT THE VSD DATA REVEALS AT A POPULATION LEVEL

| | HMO B | HMO A | HMO C |
|---|---|---|---|
| Actual HMO | Northern California Kaiser | Group Health Cooperative | Harvard Pilgrim |
| Population size | 114,966 | 15,309 | 17,547 |
| Full compliance rate at three months | 15% | 60% | 65% |
| NDD rate (%) | 1.3% | 5.7% | not done |
| Speech delay rate (%) | 2.6% | 3.9% | 4.5% |

# VACCINE SAFETY DATALINK RESULTS SHOW ASSOCIATION BETWEEN THIMEROSAL EXPOSURE AND DEVELOPMENTAL DELAY



# EXTRAPOLATING HMO POPULATION ESTIMATES TO THE TOTAL POPULATION OF THE UNITED STATES

| | |
|---|---|
| **Period of peak thimerosal exposure** | **1991-2000** |
| **Annual births per year during 1990s** | **~4 million** |
| **Total U.S. births during exposure period** | **~40 million** |
| **Rate of harm for full compliance population** | **>5%** |
| **Number of children possibly affected** | **>2 million** |

# CDC'S PRINCIPAL INVESTIGATOR COMMENTS ON INTERNAL CDC REACTIONS TO THIMEROSAL STUDY AFTER FIRST DISCUSSIONS

**"I do not wish to be the advocate of the anti-vaccine lobby and sound like being convinced [sic] that thimerosal is or was harmful, but at least I feel we should use sound scientific argumentation and not let our standards be dictated by our desire to disprove an unpleasant theory.**

**Sincerely,**

**Tom Verstraeten"**

-(***in an email to Philip Grandjean, July 14, 2000, in the aftermath of the Simpsonwood discussion***)