Letter to the Editor

The Need for Data Harmonization—A Response to Boden and Ozonoff

Boden and Ozonoff's undercount estimates in their recent Commentary rely on three assumptions for which no quantitative literature references are provided. Alternatively, we show that findings in both studies and published data indicate lower upper-bound estimates for the undercount range. Am. J. Ind. Med. 53:854–855, 2010. © *2010 Wiley-Liss, Inc.*

Boden and Ozonoff [2010], in their response to our article [Oleinick and Zaidman, 2010], clarify their earlier methodology (2008) and describe a "best-case" scenario for Minnesota undercounts of cases with days away from work (DAFW) by the Bureau of Labor Statistics (BLS) annual Survey of Occupational Injuries and Illnesses (SOII) and wage compensation cases from the state's workers' compensation database (WC) (panel 2 of their Table I). They posit that the undercounts are probably larger based on their sensitivity analyses. In contrast, we believe that panel cell values should be adjusted based on reconstructing the three report cells with available data on WC payment groups. The results strongly suggest that Boden and Ozonoff continue to underestimate the effect of the law as a confounder. The issues and approach we describe are relevant to data from most state SOII and WC databases.

The revised panel is shown below (all results rounded to the nearest hundred).

The adjustments result from the following steps:

(1) Their WC study group of 112,251 private industry study cases [Boden and Ozonoff, 2008] represented 88.1% of all cases reported through October 2002. Based on our findings, they assumed that this number included 15% WC cases paid only temporary partial disability (TPD), permanent partial disability (PPD) or by stipulation/ agreement. Since the case records for the large majority

of cases in these groups had no evidence of DAFW, we excluded them from our comparison. In response to our argument, Boden and Ozonoff dropped 13,470 cases in these three groups from cell "b," but left some 3,400 matched cases in these groups in cell "d." They characterized the reduction as a "stringent assumption" because they believed that cases in these groups were match eligible, although they were only able to match 20% of cases in this group compared to ~70% in the overall study population. We note that their assumption does not consider Ruser's [1999] finding that the national Restricted-Activity-Only case rate (no DAFW) increased from 0.25/100 FTE workers to 1.25 in 1986–1997 as the DAFW case rate dropped from 3.25 to slightly over 2.00.

Excluding the 3,400 cases from their panel column total, the SOII DAFW cases undercount estimate for Temporary and/or Permanent Total Disability (TTD/ PTD) WC cases with full days of work disability (=DAFW) is 24%.

(2) We add back 1,100 TPD/PPD/stipulation cases to cell "b" using the TTD/PTD undercount fraction since reporting rules are the same for all WC payment groups.

(3) In our study we tabulated 14,775 cases eligible for TTD/ PTD benefits on the basis of ≤3 DAFW/≤3 days TTD and now attribute 13,000 (88.1%) to private industry. Partitioning this number by the 24% undercount fraction yields 9,900 in cell "d," with the remaining 3,100 assigned to cell "b." Subtracting the 9,900 and 3,400 cases from the "d" cell matched-case total leaves 62,600 payment cases based on ≥4 DAFW. Similarly, there are 19,800 payment cases with ≥4 DAFW in "b."

(4) Boden and Ozonoff's panel 2 entry of 30,929 for cell "c" represents an arbitrary 25% reduction (10,309 cases) in response to our determination that using the date of injury to determine eligibility in the SOII database overestimates eligible cases.

†Senior Policy Analyst, Policy Development, Research and Statistics.

*Correspondence to: Arthur Oleinick, Associate Professor Emeritus, Environmental Health Sciences, School of Public Health, University of Michigan, Ann Arbor, MI 48109. E-mail: aoleinic@umich.edu

TABLE I. Undercounting Estimates, Minnesota BLS SOII and WC, Based on Integrating Results from Oleinick and Zaidman [2010], Boden and Ozonoff [2008, 2010], Literature Reports and Publicly Available Data

	Workers' compensation cases		
	No report	**Report**	**Total**
BLS SOII DAFW			
No report	"a" = 4,000−5,200	"b" = 19,800 ≥ 4 DAFW + 3,100 ≤ 3 DAFW + 1,100 TPD/PPD/stip = 24,000	28,000−29,200 (24%)
Report	"c" = 9,900−14,100 ≥ 4 DAFW + 1,600−2,200 ≤ 3 DAFW + 500−800 TPD/PPD/stip = 12,600−16,500[1]	"d" = 62,600 ≥ 4 DAFW + 9,900 ≤ 3 DAFW + 3,400 TPD/PPD/stip = 75,900	88,500−92,400
Total	16,600−21,700 (14−18%)	99,900	116,500−121,600

[1]The totals reflect reciprocal relation between DAFW group fractions.

(5) We think the reduction is still inadequate, based upon a review of the SOII data for the period 1998−2001 (provided by Mr. Zaidman). Private industry data indicate that 72,500−76,700 cases (57−60%) of DAFW cases had ≥4 DAFW (depending on how the 3−5 DAFW group is partitioned). Since 62,600 cases in this subgroup have been identified in cell "d," this leaves the balance of 9,900−14,100 in cell "c," yielding an undercount of WC cases with ≥4 DAFW compared to SOII DAFW cases of 14−18%. A similar undercount for WC cases eligible with ≤3 DAFW would imply only 1,600−2,200 such cases in "c."

In contrast, Boden and Ozonoff's [2010] estimate for "c" SOII DAFW appears to include 16,800−21,000 cases with ≤3 DAFW yielding a WC undercount percentage for this payment group of 63−68%. Their finding contrasts with the legislature's intent [Minn Stat[utes], 2009] to limit wage compensability to injuries/illnesses producing >3 days of work disability (in contrast, our study yields an estimate of 14−18% undercount for this group).

(6) We recalculate the value for "a" using the new values for "b" and "c" and assume source independence.

The above arguments do not take into account the undermatching of multi-establishment data documented for Wisconsin by Nestoriak and Pierce [2009]. Nor do they include potential matching failures due to missing matching data we documented. Consequently, we categorize the undercount estimates as upper-bound estimates.

Boden and Ozonoff [2008] have done an amazing job in estimating the number of matched cases in the two databases, but we disagree with their undercounting estimates for both the BLS SOII and WC databases. Our differences turn on how the law is treated as a confounding variable in the non-match cells.

We urge that the focus remain on developing source independence estimates without reliance on sensitivity analysis to buttress results. This is particularly true for a medical epidemiologist where there is no clear documentation for the application of huge odds ratios.

Arthur Oleinick, MD, JD, MPH,*
Associate Professor Emeritus
Environmental Health Sciences
School of Public Health
University of Michigan
Ann Arbor, Michigan

Brian Zaidman, BA[†]
Department of Labor and Industry
State of Minnesota
St. Paul, Minnesota

REFERENCES

Boden LI, Ozonoff A. 2008. Capture-recapture estimates of nonfatal workplace injuries and illnesses. Ann Epidemiol 18:500−506.

Boden LI, Ozonoff A. 2010. Researcher judgment and study design: Challenges of using administrative data. Am J Ind Med 53(1):37−41.

Minn Stat[utes]. 2009. §176.121

Nestoriak N, Pierce B. 2009. Comparing workers' compensation claims with establishments' responses to the SOII. Mon Labor Rev 132(5):57−64.

Oleinick A, Zaidman B. 2010. The law and incomplete database information as confounders in epidemiologic research on occupational injuries and illnesses. Am J Ind Med 53(1):23−36.

Ruser J. 1999. The changing composition of lost-workday injuries. Mon Labor Rev 122(6):11−17.