Data Linkage Presentation

Sujata Joshi

Project Director, IDEA-NW

Topics to discuss

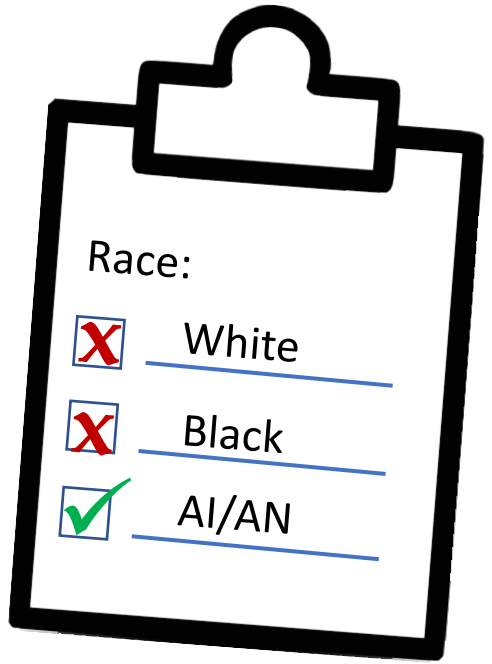
- Record linkages – purpose and approach
- WA Communicable Disease data linkages
  - Methods
  - Misclassification findings
- Next steps

What is racial misclassification?

- Racial misclassification is an incorrect recording of a person's race in a data or surveillance system

Examples:

- A person who is American Indian is coded as "white" on their death certificate
- A person who is American Indian and Black is recorded as only "Black" on their hospital visit record
- A person who is American Indian is recorded as "Hispanic White" at their doctor's office

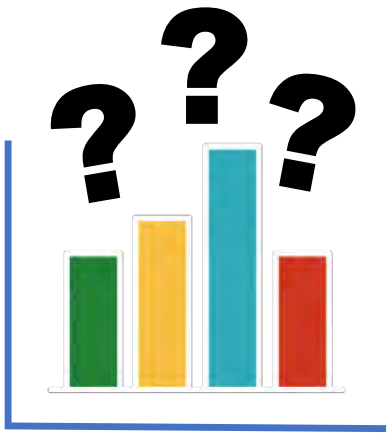


Why does misclassification matter?

When AI/AN are not classified as AI/AN in data systems, they are not counted in health reports, disease rates, or public health surveillance for AI/AN people.

This undercounting causes an **underrepresentation** of AI/AN people in the data and leads to:

- Inaccurate AI/AN health data
- Artificially lowered disease burden
- Too few AI/AN to calculate stable disease rates and trends
- Incomplete health data for public health decision-making



How common is misclassification?

Very common!

- AI/AN populations more likely to be racially misclassified than any other racial group in cancer registry datasets[1] and medical records[2]
- Published studies have found AI/AN misclassification rates ranging between 30%-70%[3,4,5]
- National death certificates study found 35,657 AI/AN (20%) were misclassified on their death certificates from 1990-2009[1]
  - These thousands of misclassified AI/AN *are not included* in AI/AN health reports, disease rates, or public health surveillance

What has NWTEC found?

Northwest Tribal Epicenter working on correcting misclassification among Northwest AI/AN since 1999

Findings:

- Depending on dataset, **7% - 65% of AI/AN misclassified**
  - Less misclassification in deaths and cancer registries
  - More in communicable disease registries, hospitalization data, trauma registries

- We've identified over **20,000 misclassified AI/AN people** in the past three years

What can we do about misclassification?

Two general ways to address misclassification:

1

Prevent it from happening

2

Fix it after it happens

What can we do about misclassification?

2

Fix it after it happens

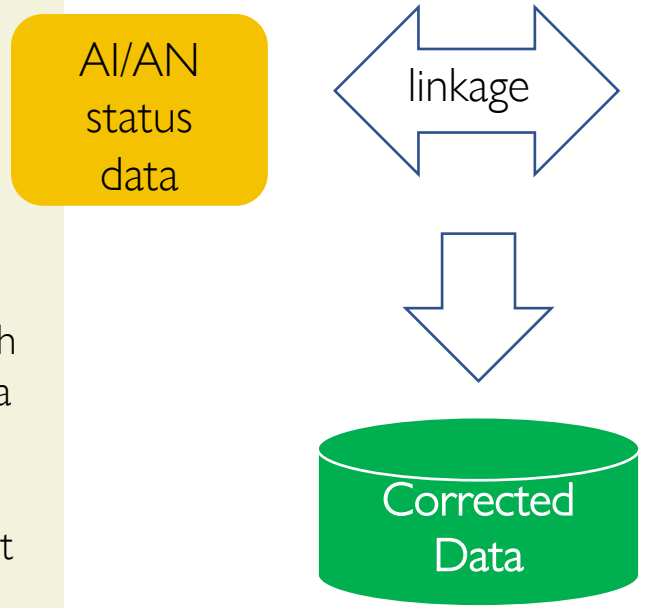
- **“Fix” race information in the data**
  - We do this with a “record linkage”
    - A record linkage is a joining of one dataset with another dataset
  1. First dataset has accurate AI/AN status information
  2. Second dataset has the health data you are hoping to analyze

2
Fix it after it happens

Linkage Overview

AI/AN status data for NWTEC "Northwest Tribal Registry"

- AI/AN seen at IHS, tribal, and urban health clinics in the Northwest
- Data obtained from the Indian Health Service with tribal approval through a resolution from NPAIHB
- Does not include any health data, just identifiers like name, address, date of birth, etc



- Death certificates
- Hospital discharge
- Cancer registries
- Syndromic Surveillance (ESSENCE)
- STD/HIV/Communicable disease
- Childhood blood lead registry
- Birth certificates
- Trauma registries
- Violent Death Registry

What do linkages accomplish?

Increased availability of accurate and complete health data for tribal communities and AI/AN

- ✓ Improved representation of AI/AN in data systems
  - ✓ Accurate AI/AN health data
  - ✓ Able to calculate more stable disease rates and trends
  - ✓ Able to provide more reliable local-level disease estimates
- ✓ Better informed public health decision-making efforts!

Washington Communicable Disease Linkages - Methods

- In 2017, we linked the Northwest Tribal Registry with 6 Washington communicable disease systems:
  - General communicable diseases
  - Hepatitis B
  - Hepatitis C
  - HIV
  - STDs
  - TB
- These linkages had the following review/approvals: NPAIHB tribal resolution, Portland Area IRB Review, Washington State IRB Review

Washington Communicable Disease Linkages - Methods

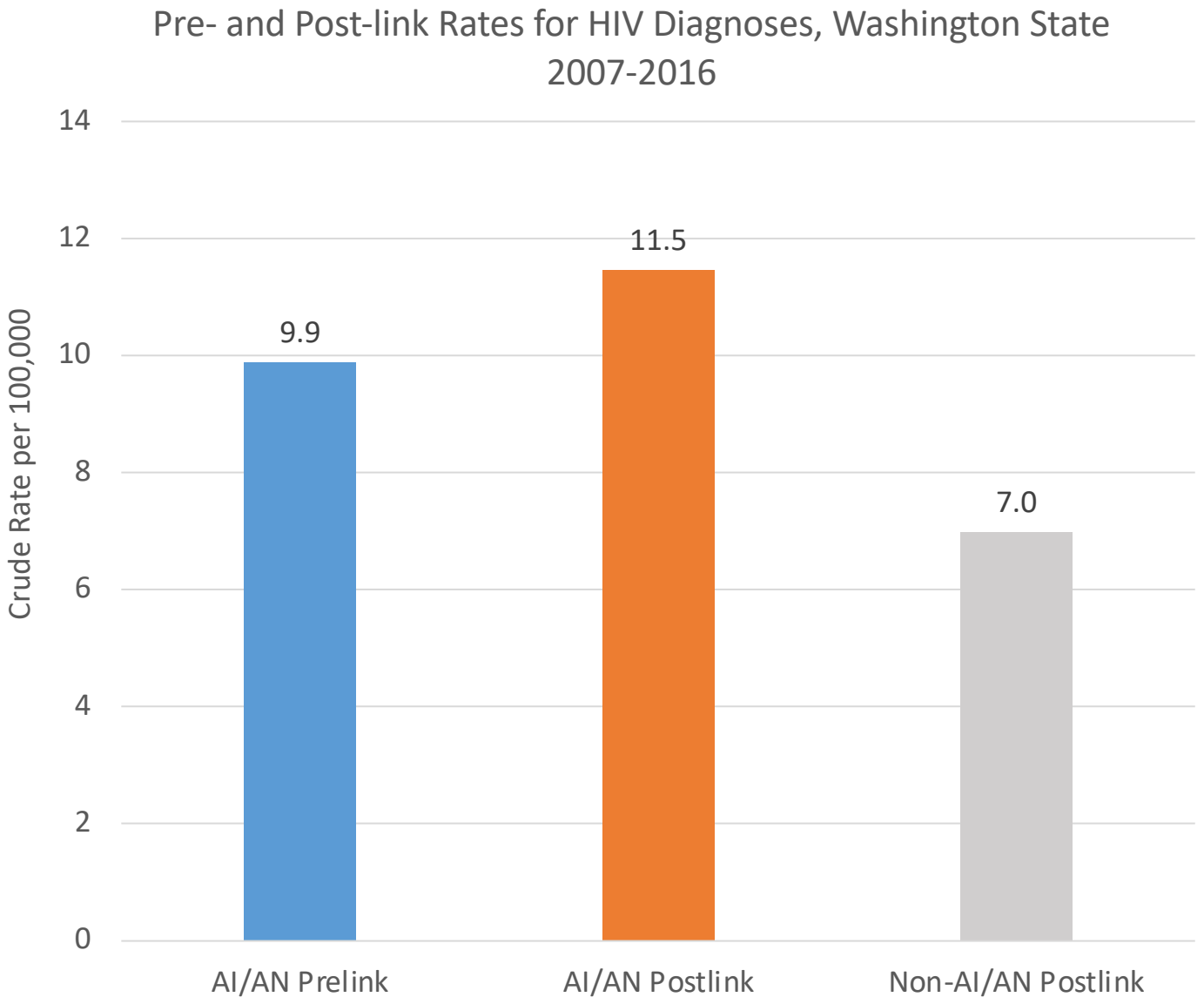
- We used a probabilistic linkage software (Link Plus) to match the WA data to the Northwest Tribal Registry on the following fields:
  - Name
  - Date of Birth
  - Sex
- After the linkage, we analyzed the data to:
  - Understand the extent of misclassification for AI/AN patients
  - Create baseline profiles of communicable diseases among AI/AN in Washington

Washington HIV Linkage

(Diagnoses from 1980-2016)

Linkage Results

- Identified 70 misclassified AI/AN HIV cases
  - Majority (78%) of misclassified records were originally coded as Non-Hispanic White
- Increased the total number of cases identified among AI/AN people by 7.8%
- Increased the rate of HIV diagnoses from 2007-2016 by 16%

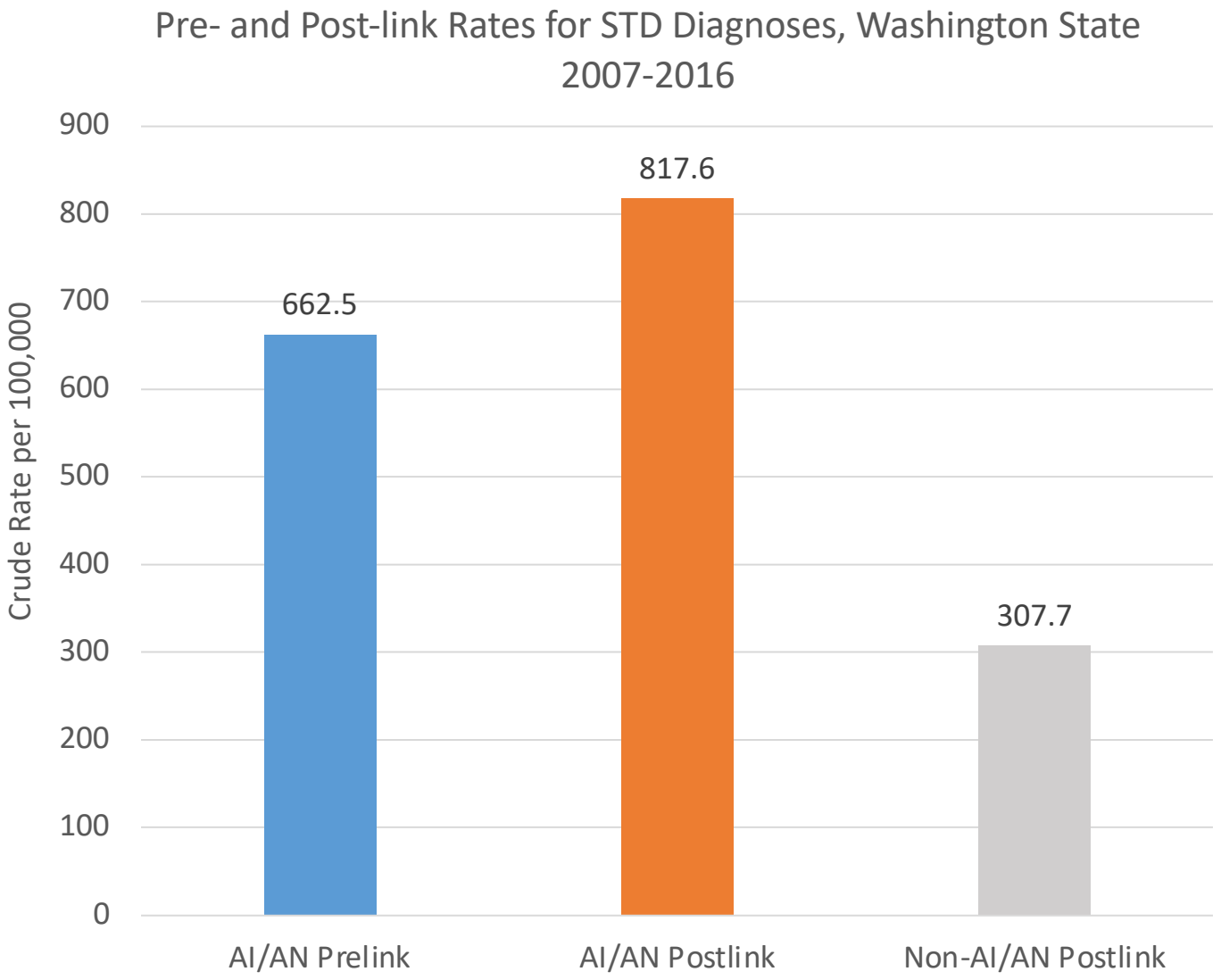


Washington STD Linkage

(Diagnoses from 2007-2016)

Linkage Results

- Identified 2,425 misclassified AI/AN STD cases
  - About 42% of misclassified STD cases were originally coded as White and 41% were originally coded as Unknown race
- Increased the total number of cases among AI/AN people by 24%
- Increased the rate of STD diagnoses from 2007-2016 by 23%

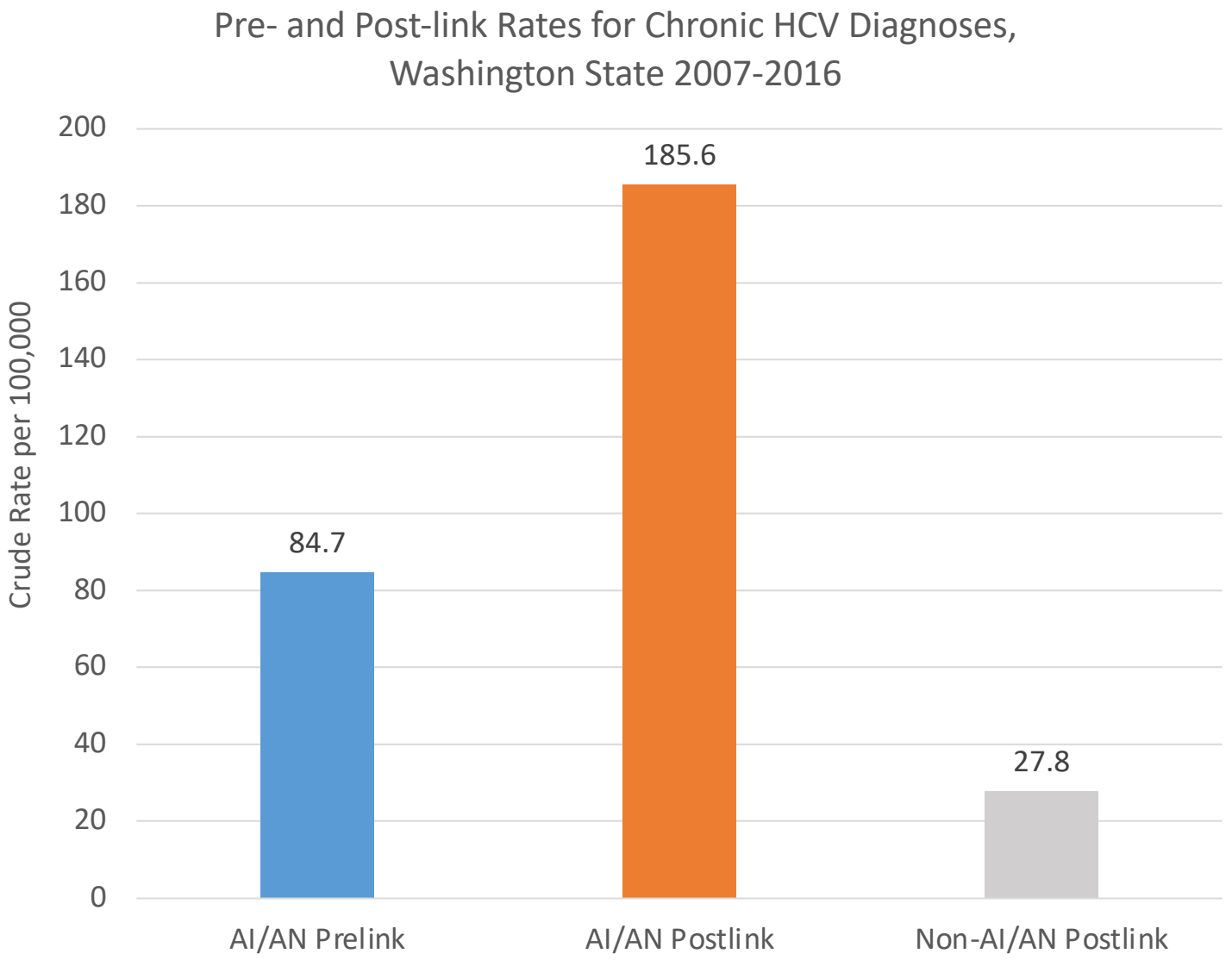


Washington Chronic HCV Linkage

(Diagnoses from 2007-2016)

Linkage Results

- Identified 1,594 misclassified AI/AN chronic HCV cases
  - Majority (86%) of misclassified records originally had missing/unknown race
- Increased the total number of cases among AI/AN people by 122%
- Increased the rate of chronic HCV diagnoses from 2007-2016 by 119%



Next Steps

- Since our 2017 linkages, Washington has consolidated their communicable disease reporting into one system (Washington Disease Reporting System)
- Next steps for NWTEC
  - Re-connect with WA DOH staff regarding access to WDRS
  - Determine interest and schedule for follow-up linkages
  - Complete linkages on a regular (every 1-2 years?) basis
  - Update disease profiles as new data become available

Questions?

Contact information:

Sujata Joshi, IDEA-NW Project Director

sjoshi@npaihb.org

Ashley Hoover, NWTEC Communicable Disease Epidemiologist

ahoover@npaihb.org

For more information on our project, please visit:

http://www.npaihb.org/idea-nw/

COVID-19 data dashboard:

http://www.npaihb.org/covid-19-data-dashboard/

References

1. Jim, M. A., Arias, E., Seneca, D. S., Hoopes, M. J., Jim, C. C., Johnson, N. J., & Wiggins, C. L. (2014). Racial misclassification of American Indians and Alaska Natives by Indian Health Service Contract Health Service Delivery Area. American journal of public health, 104 Suppl 3(Suppl 3), S295–S302. doi:10.2105/AJPH.2014.301933

2. Kressin, N.R., B. Chang, A. Hendricks, and L.E. Kazis. 2003. “Agreement between Administrative Data and Patients' Self-Reports of Race/Ethnicity.” American Journal of Public Health 93(10): 1734-1739.

3. Hoopes M., E. Vinson, and K. Lopez. 2012. “Regional Differences and Tribal Use of American Indian/Alaska Native Cancer Data in the Pacific Northwest.” Journal of Cancer Education 27(1): 73-79.

4. Johnson, J.C., A.S. Soliman, D. Tadgerson, G.E. Copeland, D.A. Seefeld, N.L. Pingatore, R. Haverkate, M. Banerjee, and M.A. Roubidoux. 2009. “Tribal Linkage and Race Data Quality for American Indians in a State Cancer Registry.” American Journal of Preventive Medicine 36(6): 549-554.

5. Boehrmer, U., Kressin, N.R., Berlowitz, D.R., Christiansen, C.L., Kazis, L.E., and Jones, J.A. 2002. “Self-Reported vs. Administrative Race/Ethnicity Data and Study Results.” American Journal of Public Health 92(9): 1471-1473.