Social, demographic, and economic correlates of food and chemical consumption measured by wastewater-based epidemiology

Phil M. Choi1,3, Benjamin Tscharkeb, Saer Samanipourb, Wayne D. Hallc, Coral E. Gartnera, Jochen F. Muellera, Kevin V. Thomasa, and Jake W. O’Briena

1Queensland Alliance for Environmental Health Science, The University of Queensland, Woolloongabba, QLD 4102, Australia; 2Norwegian Institute for Water Research, 0349 Oslo, Norway; 3Centre for Youth Substance Abuse Research, The University of Queensland, Herston, QLD 4029, Australia; and 4School of Public Health, The University of Queensland, Herston, QLD 4006, Australia

Edited by Bryan W. Brooks, Baylor University, Waco, TX, and accepted by Editorial Board Member David W. Schindler September 12, 2019 (received for review June 14, 2019)

Wastewater is a potential treasure trove of chemicals that reflects population behavior and health status. Wastewater-based epidemiology has been employed to determine population-scale consumption of chemicals, particularly illicit drugs, across different communities and over time. However, the sociodemographic or socioeconomic correlates of chemical consumption and exposure are unclear. This study explores the relationships between catchment-specific sociodemographic parameters and biomarkers in wastewater generated by the respective catchments. Domestic wastewater influent samples taken during the 2016 Australian census week were analyzed for a range of diet, drug, pharmaceutical, and lifestyle biomarkers. We present both linear and rank-order (i.e., Pearson and Spearman) correlations between loads of 42 biomarkers and census-derived metrics, index of relative socioeconomic advantage and disadvantage (IRSA), median age, and 40 socioeconomic index for area (SEIFA) descriptors. Biomarkers of caffeine, citrus, and dietary fiber consumption had strong positive correlations with IRSAD, while tramadol, atenolol, and pregabalin had strong negative correlation with IRSA. As expected, atenolol and hydrochlorothiazide correlated positively with median age. We also found specific SEIFA descriptors such as occupation and educational attainment correlating with each biomarker. Our study demonstrates that wastewater-based epidemiology can be used to study sociodemographic influences and disparities in chemical consumption.

Sociodemographic and economic determinants of population health.

The potential value of wastewater in studying the sociodemographic determinants of population health.

To date, wastewater-based epidemiology has focused on reporting drug and pharmaceutical consumption patterns by analyzing domestic wastewater. Here we explore the relationships between chemicals in wastewater and social, demographic, and economic parameters of the respective populations. We show the extent to which consumption of chemicals such as opioids and illicit drugs are associated with sociodemographics. We also examine chemicals that reflect individuals’ consumption of food components in wastewater and show that disparities in diet are associated with educational level. Our study shows that chemicals in wastewater reflect the social, demographic, and economic properties of the respective populations and highlights the potential value of wastewater in studying the sociodemographic determinants of population health.

Significance

Author contributions: P.M.C., S.S., and J.W.O. designed research; P.M.C., B.T., S.S., and J.W.O. performed research; P.M.C. and S.S. analyzed data; and P.M.C., B.T., S.S., W.D.H., C.E.G., J.F.M., K.V.T., and J.W.O. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. B.W.B. is a guest editor invited by the Editorial Board.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1To whom correspondence may be addressed. Email: p.choi@uq.edu.au.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1910242116/-/DCSupplemental.

First published October 7, 2019.

21864–21873 | PNAS | October 22, 2019 | vol. 116 | no. 43
broad, generic indicators of socioeconomic status (SES), which result in highly specific case studies rather than more generalized trends.

The use of validated and widely accepted socioeconomic metrics, such as those used by government and sociodemographic researchers, can improve the quality, interpretability, and trans-disciplinary relevance of WBE studies. The Australian Bureau of Statistics, for example, uses national census data to calculate various socioeconomic indexes for areas (SEIFA) descriptors. These are measures of income, education, employment, housing, and more. Selections of SEIFA descriptors have been used to generate broad indicators of community-level SES such as the index of relative socioeconomic advantage and disadvantage (IRSAD) (18). IRSAD has been used in various peer-reviewed studies in medical (19), environmental (20), public health (21), and behavioral (22) disciplines.

The present study examined relationships between WBE biomarkers and the population median age, average IRSAD, and 40 SEIFA descriptors (Table 1) for 22 wastewater treatment plant (WWTP) catchments in Australia (Fig. 1). Specifically, the study examined whether sociodemographic measures correlated with consumption of food components, PPCPs, and licit and illicit drugs by analyzing wastewater influent samples collected during the week of the 2016 national census. Per-capita loads of each biomarker were calculated for each WWTP catchment. Using data collected from the same census and georeferencing software, the median age (henceforth age), IRSAD, and SEIFA descriptors were calculated for the population served by each WWTP catchment. We used this to study the correlations between per-capita loads of WBE biomarkers and IRSAD, age, and SEIFA descriptors.

Results and Discussion

The present paper focuses on 42 biomarkers which were measured in wastewater above the limit of quantification with >80% frequency (SI Appendix, Table S1). The concentrations of these biomarkers in wastewater are shown in SI Appendix, Fig. S1. We report correlations between per-capita normalized loads (henceforth loads) of these substances with catchment age, catchment IRSAD, and SEIFA descriptors using Pearson (23) and Spearman rank-order correlation (24) to capture both linear and nonlinear correlations. Using the confusion matrix (25), a false positive detection rate of 0.01% corresponded to \( R = 0.35 \) (SI Appendix, Fig. S2), and we applied a more conservative significance threshold of \( R = 0.5 \).

Vitamins. We measured biomarkers of vitamins B3, E, and B6 in wastewater to assess population dietary vitamin intake. Both \( N \)-methyl-2-pyridone-5-carboxamide (2PY) and \( N \)-methyl-4-pyridone-3-carboxamide (4PY) are formed from consumption of nicotinamide, a major B3 vitamin (26, 27). \( \alpha \)-Carboxethyl hydrochroman (\( \alpha \)CEHC) is formed from the metabolism of \( \alpha \)-tocopherol (28, 29), the predominant form of vitamin E in dietary sources (30). Urinary 4-pyridoxic acid accounts for 40 to 60% of dietary vitamin B6 intake (31) and is a marker of short-term vitamin B6 intake (27, 32, 33).

We found significant linear correlations between all 4 vitamin biomarkers, particularly between the vitamin B metabolites 2PY, 4PY, and 4-pyridoxic acid (\( R = 0.555 \) to 0.895; SI Appendix, Fig. S3). This suggests that their relative intake is largely homogeneous within populations. The vitamin markers had insignificant correlation with age, but 2PY (and to a lesser extent 4PY and 4-pyridoxic acid) correlated significantly with IRSAD (\( R = 0.611 \); Fig. 2). \( \alpha \)CEHC did not reach significant correlations with any SEIFA descriptors (see Fig. 4). In contrast, all B vitamin biomarkers were most correlated with HIGHRENT (\( R = 0.534 \) to 0.662, private dwellings paying >$470 per wk in rent) and had negative correlation with descriptors of lower SES such as DISABILITYU70 (people under 70 y of age with a disability) or UNEMPLOYED (people in labor force and unemployed; Figs. 3 and 4). This agrees with the literature, as socioeconomically disadvantaged groups are less likely to meet nutritional guidelines for micro- and macronutrients (34, 35) than advantaged groups (36). Insignificant correlations between \( \alpha \)CEHC and sociodemographic measures may be due to the ubiquity of vitamin E in a range of easily accessible foods such as oils (37, 38). Interestingly, some vitamin biomarkers correlated significantly with ENGLISHPOOR (\( R_p = 0.427 \) to 0.608; Fig. 3 and SI Appendix, Fig. S6). This agrees with the trend of higher fruit and vegetable intake among Australian residents born outside of English-speaking countries (39).

Fiber and Citrus. We measured the lignin consumption biomarkers enterodiol and enterolactone and the citrus consumption biomarker proline betaine. Small and medium cohort studies have shown dose-dependent relationships between citrus consumption and urinary proline betaine excretion (40–42). Enterodiol and enterolactone form in the gut from the metabolism of the lignans matairesinol and secoisolariciresinol, which is abundant in whole grains, legumes, seeds, and fruits (43, 44). Both are urinary biomarkers of dietary fiber and polyphenol consumption from fruit and grains (45, 46) and are inversely associated with body mass index (47). We considered proline betaine, enterodiol, and enterolactone to be indicators of a healthy diet.

In our dataset, no food biomarker correlated with age, but all correlated significantly with IRSAD (Fig. 2). Correlation between the sister biomarkers enterodiol and enterolactone was strong (\( R_S = 0.886 \); SI Appendix, Fig. S4). Of all biomarkers in our study, enterodiol had the highest correlations with IRSAD (\( R_P = 0.626 \), \( R_S = 0.612 \); Fig. 2). Its greatest SEIFA correlates were OCC_MANAGER (people employed as managers, \( R_P = 0.758 \)) and NOYEAR12ORHIGHER (age >14 y and highest educational attainment year/grade 11 or lower, \( R_S = 0.720 \); Fig. 3). Likewise, proline betaine was correlated with SEIFA descriptors related to education, specifically ATUNI (age >14 y and attending university or other tertiary institution, \( R_S = 0.695 \)) and CERTIFICATE (\( R_S = -0.608 \); Fig. 4). Socioeconomically disadvantaged groups (including those in Australia) are less likely to purchase or consume grains, fruit, vegetables, and other foods high in fiber than socioeconomically advantaged groups (37–39, 48). Our results agree with these and suggest education and occupation as important factors in diet disparity.

Caffeine. Dietary caffeine intake derives primarily from coffee and tea in adult populations and soft drinks among adolescents (49). As the caffeine content of coffee (especially espresso) is generally severalfold greater than that for tea or soft drinks (49), we consider coffee consumption to be the major source of caffeine and its metabolite paraxanthine in wastewater. As expected, caffeine and paraxanthine showed strong cross-correlations, and their correlations with other biomarkers were very similar (SI Appendix, Figs. S3 and S4). Both caffeine markers correlated significantly with the markers of vitamin, citrus, and fiber consumption (\( R_P = 0.645 \) to 0.960; SI Appendix, Fig. S3). This may suggest a population-level association between consumption of caffeine and a diet rich in vitamins, citrus, and fiber, perhaps reflecting different food choices among populations with higher caffeine intake. Caffeine loads had significant linear correlation with IRSAD (Fig. 2). Additionally, caffeine and paraxanthine loads increased with HIGHRENT (\( R_P = 0.727 \), \( R_S = 0.606 \)) and decreased with CERTIFICATE (age >14 y and highest level of education is an intermediate to advanced level of vocational training, \( R_S = -0.659 \), \( R_S = -0.532 \); Figs. 3 and 4). This suggests that caffeine consumption is associated with aspects of financial capability and educational attainment.

Our results agree with an Australian dietary survey conducted in 2011 to 2012 (\( n = 6,232 \)). Habitual consumers of espresso or ground coffee lived in households with higher SES and consumed...
Table 1. SEIFA descriptors used in the present study

| Descriptors | IRSAD weighting | Definition |
|-------------|----------------|------------|
| ATUNI       | 0.36           | People aged 15 y and over who are still attending secondary school |
| DIPLOMA     | 0.5            | People aged 15 y and over whose highest level of education is a certificate III or IV qualification (intermediate or advanced vocational training) |
| DEGREE      | NA             | People aged 15 y and over whose highest level of education is a bachelor degree or higher |
| NOEDU       | −0.34          | People aged 15 y and over who have no educational attainment |
| NOYEAR12ORHIGHER | −0.85 | People aged 15 y and over whose highest level of education is year (grade) 11 or lower |
| CHILDLORLESS | −0.76         | Families with children under 15 y of age and jobless parents |
| DISABILTYU70 | −0.69         | People aged under 70 y who need assistance with core activities |
| ENGLISHPOOR | NA             | People who do not speak English well or at all |
| FEWBED      | NA             | Classifiable occupied private dwellings with one or no bedrooms |
| HIGHBED     | 0.44           | Occupied private dwellings with 4 or more bedrooms |
| GROUP       | NA             | Occupied private dwellings that are group occupied private dwellings |
| LONE        | NA             | Occupied private dwellings that are lone person occupied private dwellings |
| HIGHCAR     | −0.33          | Occupied private dwellings with 3 or more cars |
| NOCAR       | −0.33          | Occupied private dwellings with no cars |
| HIGHMORTGAGE| 0.72           | Occupied private dwellings paying more than $2,800 per mo in mortgage |
| HIGHRENT    | 0.47           | Occupied private dwellings paying more than $470 per wk in rent |
| LOWRENT     | −0.64          | Occupied private dwellings paying less than $215 per wk in rent (excluding $0 per wk) |
| INC_LOW     | −0.89          | People with stated annual household equalized income between $1 and $25,999 (approximately first and second deciles) |
| OCC_SALES_L | 0.32           | Employed people classified as Low Skill Sales |
| OCC_SERVICE_L | −0.54      | Employed people classified as Low Skill Community and Personal Service Workers |
| OCC_PROF    | 0.71           | Employed people classified as professionals |
| OCC_PROF1   | 0.71           | Employed people classified as professionals |
| OCC_SKILL1  | 0.65           | Families that are one-parent families with dependent offspring only |
| INC_HIGH    | 0.83           | People with stated annual household equalized income greater than $78,000 (approximately ninth and tenth deciles) |
| MORTGAGE    | NA             | Occupied private dwellings owning the dwelling they occupy with a mortgage |
| ownership   | NA             | Occupied private dwellings owning the dwelling they occupy without a mortgage |
| NOONE_PARENT| −0.65          | Families that are one-parent families with dependent offspring only |
| UNINCORP    | NA             | Owner of an unincorporated enterprise |
| UNEMPLOYED1 | NA             | People aged 15 y and over who are unemployed |
| UNEMPLOYED2 | NA             | People (in the labor force) who are unemployed |
| OCC_SALES_L | −0.32          | Employed people classified as Low Skill Sales |
| OVERCROWD   | −0.33          | Occupied private dwellings requiring one or more extra bedrooms (based on Canadian National Occupancy Standard) |

NA: not applicable. Details regarding variable composition can be found elsewhere (18).

more dietary caffeine than those who consumed mix coffee or instant coffee, or those not consuming coffee (50). Espresso and ground coffee consumers were also more likely to have a bachelor’s degree or higher than other groups (50), which is analogous to our findings with the SEIFA descriptors. Our findings also support cohort studies which find higher caffeine consumption among higher SES adolescents (51) and adults (52). Although this trend may be reversed for specific demographics at risk for caffeine abuse (53), our results suggest the prevailing population-scale trend is a mild correlation between caffeine consumption and IRASD. We suggest that increased caffeine consumption in socioeconomically advantaged groups may reflect 1) greater financial freedom to indulge in caffeinated beverages (i.e., coffee) and/or 2) cultural institutionalization of regular coffee drinking among advantaged and/or educated populations.

Opioids. We analyzed the opioids methadone and its metabolite 2-ethylidene-1,5-dimethyl-3,3-di phenyl pyrrolidine (EDDP), codeine, morphine, oxycodone and its metabolite noroxycodone, and tramadol. Codeine was the only opioid available over the counter (OTC) during the time of sampling. Loads of codeine, oxycodone, and noroxycodone were significantly correlated with each other (RS = 0.516 to 0.724; SI Appendix, Fig. S4). This reflects catchment-level polydrug use and the coexcretion of oxycodone and noroxycodone following oxycodone consumption (54, 55). Similarly, methadone and EDDP correlated (RS = 0.948) as expected (SI Appendix, Fig. S3).
Morphine was the only opioid significantly correlating with age ($R_S = 0.543$; Fig. 2), which matches its primary use in chronic pain which increases with age (56, 57). It should be noted that morphine residues can also be present from the metabolism of other opioids such as codeine and heroin (58). Significant correlations with IRSAD were only seen for tramadol ($R_S = -0.701$) and codeine ($R_S = 0.539$; Fig. 2), suggesting that a broad SES indicator does not explain methadone, morphine, and oxycodone consumption at a population scale. However, strength of correlations was higher with specific socioeconomic descriptors. SEIFA correlates of all opioids studied showed moderate to strong correlations with descriptors of socioeconomic disadvantage, and strong inverse correlations with SEIFA measures of socioeconomic advantage (Fig. 4). As illustrated in Fig. 3, methadone loads were slightly lower in catchments with fewer ENGLISHPOOR (individuals with poor English-speaking ability) and more residents with limited educational attainment. Codeine loads were lower in catchments with fewer HIGHMORTGAGE (private dwellings paying more than $2,800 per mo in mortgage) and were generally higher in catchments with more ONEPARENT (one-parent families with dependent offspring only). Oxycodone loads increased with the proportion of DISABILITYU70 and decreased with increasing household income. Tramadol loads were lowest in populations with greater percentage of skilled workers or a low percentage of laborers. Overall, these results are in line with the demographic correlates of opioid use. Opioid users are more likely to be taking other opioids (54) and have other health and lifestyle complications (59). In Australia, laborers and groups with limited education are more likely to use opioids (60). Likewise, in other countries opioid use is associated with limited educational achievement, lower household income, and other substance abuse behavior (61, 62), and these match our findings regarding the SEIFA variables. However, methadone, codeine, (nor)oxycodone, and tramadol use were not associated with catchment age, which can be considered a proxy for algesia.

**Antidepressants.** We measured the antidepressants citalopram, desvenlafaxine, its parent compound venlafaxine, mirtazapine, and amitriptyline, all prescription pharmaceuticals with little difference in price during the time of sampling. We considered antidepressants as a proxy for psychological distress. All antidepressants tested had no significant association with age (SI Appendix, Figs. S3 and S4). All antidepressants tested except citalopram had significant inverse correlations with IRSAD ($R_S = -0.529$ to $-0.701$; Fig. 2). Interestingly, these antidepressants seemed to diverge in terms of SEIFA correlates. Desvenlafaxine had strongest positive correlations with OCC_LABOR (people employed as laborers, $R_P = 0.629$ and 0.549, respectively; Fig. 4) and significant correlations with similar educational and occupation descriptors (SI Appendix, Figs. S6 and S7). Amitriptyline was most influenced by educational descriptors; NOYEAR12ORHIGHER ($R_P = 0.770$) and DEGREE (age ≥14 y and highest educational attainment a bachelor degree or higher, $R_S = -0.674$; Fig. 2). In contrast, citalopram had stronger correlations with social descriptors such as LONE (private dwellings with one resident), SEPDIVORCED (age ≥14 y and separated or divorced), and NONET (occupied private dwellings with no internet) than with education and occupation descriptors (Fig. 4 and SI Appendix, Figs. S6 and S7), hinting at a difference in demographics between des/venlafaxine, amitriptyline, and citalopram consumers. In Australia between 2003 and 2005 there were significant associations between antidepressant use and lower SES among those aged ≥30 y but a statistically insignificant trend in those aged 20 to 29 y and a reverse association in those under 19 y (63). The lack of a consistent trend encompassing all age groups may explain why correlations were not significant when correlated with IRSAD but significant with SEIFA descriptors.

We found significant correlations between antidepressants and methadone, codeine, morphine, tramadol, and pregabalin (SI Appendix, Figs. S3 and S4). This may reflect individual and/or population-scale polydrug use; patients prescribed opioids are also often prescribed antidepressants (64), and Australian pregabalin users often coingest opioids (65).

**Anticonvulsants.** We measured the anticonvulsants carbamazepine, gabapentin, and pregabalin. Carbamazepine and gabapentin had insignificant correlation with increasing age, and carbamazepine also had insignificant correlations with decreasing IRSAD (SI Appendix, Figs. S3 and S4). Pregabalin was significantly correlated with increasing age ($R_P = 0.510$) and decreasing IRSAD ($R_S = -0.731$; Fig. 2). Pregabalin correlated significantly with opioids and antidepressants, and this was true to a lesser extent for carbamazepine and gabapentin (SI Appendix, Figs. S3 and S4). These findings are expected as anticonvulsants are also used to treat neuropathic pain (66). The different anticonvulsants seem to be associated with age, IRSAD, and other biomarkers of distress to differing extents (SI Appendix, Figs. S3 and S4), suggesting consumption of each is driven by different sociodemographic drivers. Fig. 3 shows pregabalin loads were lowest in catchments with greater proportion of the population with INC_HIGH (household income in highest 2 deciles, $R_S = -0.767$) and in catchments with fewer DISABILITYU70 ($R_S = 0.745$). The second most positive and negative SEIFA correlates of pregabalin were INC_LOW (household income in lowest 2 deciles, $R_S = 0.688$) and HIGHMORTGAGE ($R_S = -0.748$), highlighting income as an important factor for pregabalin use (SI Appendix, Figs. S6 and S7).

Pregabalin and gabapentin abuse is of increasing public health concern (67); pregabalin use in particular has also been linked to increased suicide risk in Australia. The links between lower SES and higher pregabalin use, as well as its correlation with opioids, antidepressants, and atenolol are in agreement with trends observed in Greece during a worsening social and financial crisis (16). Therefore, opioids, antidepressants, anticonvulsants, and atenolol may be considered proxies of socioeconomic distress.

**Illicit Drugs.** Methamphetamine is an illicit drug, and around 4% of a dose of methamphetamine is excreted as amphetamine (68). This seems to be reflected in the significant correlations...
between the 2 biomarkers (SI Appendix, Figs. S3 and S4). Amphetamine can be sourced from both illicit and legal sources, although licit usage is negligible at a population scale (SI Appendix, Supplementary Information Text 1). Amphetamine and methamphetamine had insignificant correlations with IRSAD and age (Fig. 2) but both had moderate inverse correlations with SEIFA descriptors of socioeconomic disadvantage (Fig. 4). For example, methamphetamine consumption was relatively lower in catchments with fewer OCC_SKILL5 (low-skilled workers, RP = 0.625; Fig. 3). Conversely, catchments with more than 15% of the population with DIPLOMA (highest educational attainment is a diploma) had relatively low methamphetamine consumption (Fig. 3). The influence of SES on methamphetamine use is multifaceted and difficult to segregate by broad socioeconomic indices such as IRSAD and age. However, the SEIFA descriptors best correlating with their use are similar to demographic characteristics of known meth/amphetamine users (69, 70).

Alcohol. Ethyl sulfate, a biomarker of ethanol (alcohol) consumption (71), was not correlated with age (RP = −0.081) and had insignificant correlation with IRSAD (RP = 0.449; Fig. 2). This result agrees with an Australian dietary recall study (n = 9,345) conducted in 2011 to 2012 which found insignificant difference between alcohol intake among SES quintiles and most age groups (39). However, ethyl sulfate was significantly correlated with HIGHRENT (RP = 0.716; Figs. 2 and 4) and similar indicators of socioeconomic advantage such as INC_HIGH (RP = 0.601) and OCC_MANAGER (RP = 0.537; SI Appendix, Figs. S6), indicating higher alcohol consumption in these socio-demographically advantaged groups.

Tobacco. The tobacco consumption markers hydroxycotinine, cotinine, and nicotine all strongly correlate with each other (SI Appendix, Figs. S3 and S4). This is expected, as hydroxycotinine and cotinine are metabolites of nicotine and their excretion is relatively consistent between individuals in urine (72). No tobacco biomarkers significantly associated with catchment age or IRSAD (Fig. 2). However, with the exclusion of an outlier catchment, nicotine had significant negative correlations with IRSAD (RP = −0.504; SI Appendix, Supplementary Information Text 2). This suggests that correlation between IRSAD and tobacco consumption is not robust at a population scale. The Australian National Drug Strategy Household Survey of 2016 found smoking prevalence in the
lowest and highest IRSAD quintiles to be 20.0% and 9.3%, respectively (73). Such a trend may have been difficult to detect in free-living populations as they are not strictly segregated by SES. Figs. 3 and 4 show hydroxycotinine and cotinine correlated significantly with NONET (RS = 0.512 to 0.527). Inverse correlations with MORTGAGE (dwellings occupied with a mortgage) or ATSCHOOL (individuals aged 15 y or older attending secondary school) were insignificant (RS = −0.472 to −0.484). This suggests that tobacco use may be higher among the socially isolated, and lower among the middle to upper-middle class. These findings agree with studies in Australian and other populations that found inverse correlations between nicotine use and various proxies of SES (21, 74).

Nonsteroidal Antiinflammatory Drugs. We studied the OTC NSAIDs ibuprofen and naproxen, which differ in their duration of efficacy. The relatively short-acting ibuprofen is used to treat acute pain, whereas naproxen is longer-acting and better suited to...
Antihistamines. Fexofenadine and cetirizine are both OTC antihistamines used to treat allergy symptoms and can be considered biomarkers of population histamine burden in a WBE context (76). Significant correlations between the antihistamines and hydroxyzine and cotinine (SI Appendix, Figs. S3 and S4) may reflect the greater prevalence of asthma among individuals exposed to tobacco smoke (77, 78). Neither antihista-

mine was associated with age or IRSAD in our dataset (|R| = 0.05 to 0.384; Fig. 2). This is expected as general socioeconomic indices are not a consistent correlate of allergic disease incidence despite higher allergic disease incidence in urbanized populations (77, 79, 80). Rather, environmental factors such as pollen, dust, and dampness may be more important drivers of allergic disease (77, 81).

Antihypertensives. Atenolol was significantly correlated with increasing age (R_P = 0.660) and decreasing IRSAD (R_S = −0.657; Fig. 2). This is anticipated since atenolol is a commonly prescribed (82) beta blocker for hypertension (83). Atenolol loads decreased with HIGHMORTGAGE (R_S = −0.763) and increased with NONET (R_S = 0.803; Fig. 3). This confirms the expectation of a higher prevalence of hypertension in older and socioeconomically disadvantaged individuals (84, 85). Atenolol also had significant correlations with several opioids and antidepressants (SI Appendix, Figs. S3 and S4). This is consistent with greater use of analgesics among older individuals and supports multiple cohort studies which show depression increases the risk of hypertension (86).

Hydrochlorothiazide is an OTC diuretic often used as a first-line treatment for hypertension. Like atenolol, hydrochlorothiazide correlated with age, but it contrasts with atenolol in its lack of significant correlation with IRSAD (Fig. 2) and lower correlations with SEIFA descriptors (Fig. 4). This may reflect its use as a first-line of treatment and therefore a more uniform use throughout different sociodemographic groups.

Antibiotics. We measured markers of the antibiotics cephalexin (a cephalosporin and broad-spectrum antibiotic), sulfamethoxazole (a sulphonamide), and trimethoprim. Significant correlation between sulfamethoxazole and trimethoprim (SI Appendix, Figs. S3 and S4) reflects the fact that the 2 drugs are commonly prescribed in combination for bacterial infections. No antibiotics had significant correlation with age or IRSAD (|R| = 0.006 to 0.280; Fig. 2). In particular, sulfamethoxazole and trimethoprim had insignificant correlation with SEIFA variables (|R| = 0.161 to 0.300; Figs. 3 and 4). Our results suggest that sociodemographic factors have no major bearing on antibiotic use (particularly sulfamethoxazole and trimethoprim) at a population scale. The literature on antibiotic use and sociodemographics is mixed; European studies have shown a negative linear relationship (87, 88), including use of cephalosporins, sulfonamides, and trimethoprim (89). However, other studies in Europe and the United States highlight physician density and renumeration as major predictors of antibiotic use (88, 90).

Sweeteners. Artificial sweeteners are found in beverages, condiments, breads, cereals, dairy products, and pharmaceuticals (91). We measured levels of acesulfame, saccharin, and sucralose. The artificial sweeteners had insignificant correlations with age or IRSAD (Fig. 2), with the strongest correlation being between sucralose and age (R_S = −0.445). Acesulfame in particular had negligible correlations with age (R_P = −0.164) and IRSAD (R_S = −0.065). Similarly, correlations between sucralose and SEIFA descriptors were insignificant (Figs. 3 and 4). Of the artificial sweeteners, only saccharine reached significant correlation with a SEIFA descriptor, OWNING (dwellings owned without a mortgage, R_P = 0.538; Fig. 4). Our results suggest mildly lower saccharine consumption among populations which are older but not socioeconomically disadvantaged. Results regarding acesulfame support an earlier WBE study that found per-capita loads to have a strong linear relationship (R^2 = 0.995) with catchment population size of 10 WWTPs spread out over 4 Australian states and territories (92).

Assumptions and Limitations. This study has certain methodological limitations. As the study deals with correlations between sociodemographic outcomes and biomarkers, we assume that the presence of a biomarker of consumption in the wastewater solely treating chronic pain (75). Insignificant correlation (R_P = 0.425) between ibuprofen and naproxen (SI Appendix, Figs. S3) suggests limited relationship between the use of these NSAIDs. Ibuprofen loads were not correlated with age or IRSAD (Fig. 2), and correlations with SEIFA descriptors were also insignificant (Fig. 4), suggesting relatively uniform use across the sociodemographic spectrum. In contrast, naproxen approached significant correlation with age (R_P = 0.425) but not IRSAD (Fig. 2 and SI Appendix, Figs. S3 and S4) and was inversely correlated with ATSCHOOL (R_P = −0.619, people aged 15 y and over attending secondary school; Fig. 4), suggesting lower use in populations with younger people. Overall, our results suggest that naproxen use is slightly lower in younger catchments, which agrees with its application for chronic pain (75).

Fig. 4. Linear or first-order correlation coefficients (R) of normalized biomarker loads with SEIFA descriptors, whose definitions are supplied in Table 1. Correlation type is designated P (Pearson) or S (Spearman). The significance cutoff was |R| = 0.5.
reflects consumption by individuals. This is reasonable for biomarkers such as antidepressants. It may be less accurate for illicit drugs, which potentially may be discarded via the wastewater stream, and for proline betaine, an osmoregulator in citrus flesh, which may enter the sewer system via food scraps or industrial processing of citrus. However, the association of proline betaine with IRSAD (Fig. 2) and ATUNI (Fig. 4) implies higher loads from metropolitan catchments less likely to have wastewater input from food processing industries. This suggests that industrial processing were not a major issue in our study.

In-sewer transformation can affect some biomarkers. This means that the load that enters the sewer system may differ from the load sampled (93). In planning this study, we avoided measuring biomarkers known to transform rapidly in sewer systems such as paracetamol (94). The extent of in-sewer transformation is catchment-specific and is not known for all chemicals tested in this study, in some cases due to the novel nature of their application.

The loads of a biomarker in wastewater do not give insight into the prevalence of its excretion in the population sampled. For example, if different catchments have the same load of a biomarker, we cannot determine whether these communities have the same number of people exposed at the same level or a fewer number of people exposed at a higher level. This is a limitation of the method but it is an advantage from an ethical viewpoint because it ensures the anonymity of individuals in the catchment area. Hall et al. (95) provide suggestions on how WBE studies should be reported to avoid any adverse impacts on residents of wastewater catchment areas.

Our wastewater samples came from 21.1% of the Australian population. Our study was limited to catchments from where samples were available and so is not a randomly selected subset of the Australian population. Nevertheless, the selected locations ensured that the proportion of the population in urban catchments as compared to rural catchments in our dataset approximated to the proportions in the overall Australian population.

The census-derived data used to calculate median age, IRSAD, and SEIFA for the catchment areas were made available in georeferenced mesh blocks. When the boundary of a WWTP intersected mesh blocks, the proportion of area inside the WWTP catchment was used to scale the mesh-block calculation. This calculation assumed a uniform distribution of population inside each mesh block, which is unlikely in most cases. Nevertheless, these uncertainties are expected to be minimal as each mesh block contains a relatively small number of people (96).

In terms of the correlation analysis, we used Pearson linear correlation (i.e., parametric) and Spearman rank-order correlation (i.e., nonparametric). The parametric approach is much more sensitive to trends in the data but assumes a normal distribution for the data. On the other hand, nonparametric methods (i.e., Spearman correlation) are less sensitive and do not assume any data distribution. We combined the 2 methods rather than transforming the data and using the parametric approach. Additionally, our approach can only capture the direct bivariate correlation between different descriptors; more sophisticated multivariate methods would be able to isolate the associations with a combination of descriptors. Finally, the correlation coefficient threshold for statistical significance was defined using the confusion matrix approach (i.e., bootstrapping). This method only provides a probability of false detection for each correlation coefficient value, which has its own limitations (97).

**Conclusion**

This WBE study examined the relationship between WBE biomarkers and sociodemographic descriptors. In general, vitamin, citrus, and fiber consumption was associated with IRSAD and educational descriptors, highlighting socioeconomic, educational, and occupational status as important factors in diet quality. Conversely, lower IRSAD was associated with higher loads of tramadol, desvenlafaxine, mirtazapine, pregabalin, and atenolol. This was consistent with higher pharmaceutical use in low SES groups. Increasing population age was associated with morphine, citrulon, atenolol, and hydrochlorothiazide, which may be expected from their increase use among the elderly population. Although WBE has primarily been used for measuring drug consumption, our results demonstrate that it can be used to identify sociodemographic patterns or disparities which associate with consumption of specific chemicals or food components.

**Materials and Methods**

Wastewater samples analyzed in this study were collected from Australian WWTPs as 7 consecutive daily samples taken around the time of the 2016 Australian census (9 August 2016) (96). Each sample was sampled flow- or time-proportionally. The WWTP catchments were carefully selected to maintain a ratio of urban and rural catchments representative of the greater Australian population (Fig. 1). Twenty-two WWTPs from 6 states and territories were sampled that incorporated 21.1% of the Australian population. Catchment populations ranged from 1,000 to 2.5 million (approximated for anonymity). Five to seven consecutive 24-h composite samples, one of which was collected on census day, were collected from the influent of each WWTP. Each flow- or time-proportional sample was adjusted to pH 2 with 2 M HCl and frozen immediately after sampling. Samples were transported to the laboratory on ice and stored at −20 °C until analysis. Each sample was filtered through a 0.45 μm filter (Agilent). One milliliter aliquots were spiked with internal standard to 10 μg/l for analysis. Detailed sample processing procedures are outlined elsewhere (96).

By sampling wastewater at the time of the census, wastewater samples could be matched with sociodemographic data collected in the census. Catchment populations were determined by overlaying georeferenced WWTP catchment maps onto mesh blocks. Details of this methodology can be found elsewhere (96). In brief, the population counts were summed within each catchment area and allocated on a proportional-area basis for mesh-block units which intersected the catchment boundary. When the boundary of a WWTP intersected mesh blocks, the proportion of area inside the WWTP catchment was used to scale the mesh-block calculation. This calculation assumed uniform distribution of population inside each mesh block. Nevertheless, uncertainties arising from these artifacts are expected to be minimal as 99.9% of mesh blocks contain <500 people (96).

IRSAD is devised by the Australian Bureau of Statistics using a selection of SEIFA descriptors to summarize the economic and social conditions of people and households in the area by using descriptors that reflect relative socioeconomic advantage and disadvantage. A high IRSAD score indicates socioeconomic advantage, and a low IRSAD score indicates socioeconomic disadvantage. A full list of the SEIFA descriptors including the loadings of descriptors used to form IRSAD can be found in Table 1. The technical paper describing the construction of IRSAD is available online (18). In our study, average IRSAD was calculated specifically for each catchment in the same way as age, but using Statistical Area level 1 (SA1) resolution (the second-smallest census unit, typical population 200 to 800 people per unit). SEIFA descriptors were calculated likewise using the SA1 resolution.

SI Appendix, Table S1 lists the Chemical Abstracts Service number and standard manufacturer of each biomarker. All standards were of analytical grade. Reagents used for the analysis were analytical-grade methanol from Merck Pty Ltd and acetic acid from Sigma; HCl was from Merck Pty Ltd. Water was purified using a MilliQ system (0.22-μm-filtered, 18.2 MΩ/cm; Millipore). Biomarkers were quantified by isotope dilution or internal standards in 2 parallel liquid chromatography tandem mass spectrometry methods. Method A was used to measure mostly illicit and licit drugs and is documented elsewhere (98). Briefly, a Shimadzu Nexera UHPLC system housing a Phenomenex Kinetex 2.6 μm Biphelopy (100 Å, 50 × 2.1-mm) column was connected to an AB SCIEX QTRAP mass spectrometer. For both methods, an 8-point calibration series was used to quantify chemicals using isotope dilution methods. A typical run featured blank, QTRAP, and QC sample every 6 to 8 samples. Repeatability and validation details for method B are provided in SI Appendix, Tables S3–S5.

Biomarker loads in each sample were converted to milligrams per 24 h per 1,000 people using census-derived populations for each WWTP catchment. We did not perform any other data pretreatment in order to...
minimize the potential of false correlations caused by data pretreatment processes.

ACKNOWLEDGMENTS. We thank the numerous wastewater treatment plant and Queensland Alliance for Environmental Health Science personnel involved in sampling wastewater from each catchment and Dr. Katherine Langford for proofreading and helpful comments. The Queensland Alliance for Environmental Health Sciences, The University of Queensland, gratefully acknowledges financial support of the Queensland Government. This project was supported by an Australian Research Council Linkage Project (LP150100364), a Linkage Infrastructure, Equipment and Facilities grant (LE110100032), and a University of Queensland Major Equipment and Infrastructure grant.

1. E. Zuccato, C. Chiabrando, S. Castiglioni, R. Bagnati, R. Fanelli, Estimating community drug abuse by wastewater analysis. Environ. Health Perspect. 116, 1027–1032 (2008).
2. K. V. Thomas et al., Comparing illicit drug use in 19 European cities through sewage analysis. Sci. Total Environ. 432, 432–439 (2012).
3. J. A. Baz-Lomba et al., Comparison of pharmaceutical, illicit drug, alcohol, nicotine and caffeine levels in wastewater with sale, seizure and consumption data for 8 European cities. BMC Public Health 16, 1035 (2016).
4. G. Golovko, V. Kumar, G. Fedorova, T. Randak, R. Grabic, Seasonal changes in antidepressant and psychotropic drugs, antihistamines and lipid regulators in a wastewater treatment plant. Chemosphere 111, 418–426 (2014).
5. W. Wang, K. Kannan, Mass loading and emission of benzophenone-3 (BP-3) and its derivatives in wastewater treatment plants in New York State, USA. Sci. Total Environ. 579, 1316–1322 (2017).
6. B. Kszpircz-Hordern, R. M. Dinsdale, A. J. Guvey, The occurrence of pharmaceuticals, personal care products, endocrine disruptors and illicit drugs in surface water in South Wales, UK. Water Res. 42, 3498–3518 (2008).
7. J. W. O’Brien et al., Wastewater analysis of Census day samples to investigate per capita input of organophosphorous flame retardants and plasticizers into wastewater. Chemosphere 138, 328–334 (2015).
8. D. B. Raker, B. Kszpircz-Hordern, Multi-residue determination of the sorption of illicit drugs and pharmaceuticals to wastewater suspended particulate matter using pressure-saturated liquid extraction, solid phase extraction and liquid chromatography coupled with tandem mass spectrometry. J. Chromatogr. A. 1218, 7901–7913 (2011).
9. K. V. Thomas, M. J. Reid, What else can the analysis of sewage for urinary biomarkers reveal about communities? Environ. Sci. Technol. 45, 7611–7612 (2011).
10. A. K. Venketesan, J. Chen, E. Driver, A. Guoghar, R. U. Halden, “Assessing the potential to monitor plant-based diet trends in communities using a wastewater-based epidemiology approach” in Wastewater-Based Epidemiology: Estimation of Community Consumption of Drugs and Diets. B. Subedi, Ed. (American Chemical Society, 2019), pp. 1319, pp. 187–198.
11. R. S. Mackie et al., Trends in nicotine consumption between 2010 and 2017 in an Australian city using the wastewater-based epidemiology approach. Environ. Int. 125, 184–190 (2019).
12. A. J. Lloyd et al., Using wastewater-based analysis to monitor the effects of legal retail sale of cannabis in Washington State, USA. Addiction 114, 1582–1590 (2019).
13. C. Harman, M. Reid, K. V. Thomas, In situ calibration of a passive sampling device for selected illicit drugs and their metabolites in wastewater, and subsequent year-long assessment of community drug usage. Environ. Sci. Technol. 45, 5676–5682 (2011).
14. D. Phung et al., Can wastewater-based epidemiology be used to evaluate the health impact of temperature? An exploratory study in an Australian population. Environ. Res. 156, 113–119 (2017).
15. C. G. Daughton, Monitoring wastewater for assessing community health: Sewage Chemical-Information Mining (SCM). Sci. Total Environ. 619–620, 748–764 (2018).
16. N. S. Thomaidis et al., Reflection of socioeconomic changes in wastewater: Licit and illicit drug use patterns. Environ. Sci. Technol. 50, 10065–10072 (2016).
17. Y. Zhang et al., Wastewater-based epidemiology in Beijing, China: Prevalence of antibiotic use in flu season and association of pharmaceuticals and personal care products with socioeconomic characteristics. Environ. Int. 125, 152–160 (2019).
18. Australian Bureau of Statistics, “Census of population and housing: Socio-economic indexes for areas (SEIFA), Australia, 2016” (Australian Bureau of Statistics, Commonwealth of Australia, 2018).
19. D. A. González-Chica, S. Vanlent, E. Hoon, N. Stocks, Epidemiology of arthritis, chronic back pain, gout, osteoporosis, spondyloarthropathies and rheumatoid arthritis among 1.5 million patients in Australian general practice: NPS MedicineWise MedicareInsight dataset. BMC Musculoskelet. Disord. 19, 20 (2018).
20. D. Kendall, N. S. G. Williams, K. J. H. Williams, Drivers of diversity and tree cover in gardens, parks and streetscapes in an Australian city. Urban For. Urban Green. 15, 257–265 (2012).
21. L. J. Burt et al., G. Pereira, N. Middleton, S. Foster, Socioeconomic area disparities in tuberculosis retail outlet density: A Western Australian analysis. Med. J. Aust. 198, 489–491 (2013).
22. A. L. Gamble et al., Adolescent sleep patterns and night-time technology use: Results of the Australian Broadcasting Corporation’s Big Sleep Survey. PLoS One 9, e11700 (2014).
23. L. I. Lin, A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255–268 (1989).
24. C. Spearman, The proof and measurement of association between two things. Am. J. Psychol. 15, 72–101 (1904).
25. R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection” in Proceedings of the 14th International Joint Conference on Artifi- cial Intelligence (Morgan Kaufmann, San Francisco, 1995), vol. 2, pp. 1137–1145.
26. T. T. T. Fukuwari, S. Sasaki, K. Shibata, Twenty-four-hour urinary water-soluble vitamin levels correlate with their intakes in free-living Japanese schoolchildren. Public Health Nutr. 14, 327–333 (2011).
