BMJ Open

Monte Carlo simulation assess variability and uncertainty of tobacco consumption in a city by sewage epidemiology

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010583
Article Type:	Research
Date Submitted by the Author:	19-Nov-2015
Complete List of Authors:	Wang, De-Gao; Dalian Maritime University, School of Environmental Science and Technology Dong, Qian-Qian; Dalian Maritime University, School of Environmental Science and Technology Du, Juan; Dalian Maritime University, School of Environmental Science and Technology Yang, Shuo ; Dalian Maritime University, School of Environmental Science and Technology Zhang, Yun-Jie ; Dalian Maritime University, Department of Mathematics Na, Guang-Shui ; National Marine Environmental Monitoring Center, Ferguson, Stuart ; University of Tasmania, Faculty of Health Science, School of Medicine Wang, Zhuang ; Nanjing University of Information Science and Technology, Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), School of Environmental Science and Engineering Zheng, Tong ; Harbin Institute of Technology, State Key Laboratory of Urban Water Resource and Environment,
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Addiction, Epidemiology, Smoking and tobacco
Keywords:	wastewater analysis, Monte Carlo simulation, cotinine, trans-3'-hydroxycotinine, tobacco use

Monte Carlo simulation assess variability and uncertainty of tobacco consumption in a city by sewage epidemiology

De-Gao Wang[1†*], Qian-Qian Dong[1], Juan Du[1], Shuo Yang[1], Yun-Jie Zhang[2], Guang-Shui Na[3], Stuart G. Ferguson[4], Zhuang Wang[5], Tong Zheng[6]

[1]School of Environmental Science and Technology, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[2]Department of Mathematics, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[3]National Marine Environmental Monitoring Center, Dalian, Liaoning, China, 116023

[4]Faculty of Health Science, School of Medicine, University of Tasmania Private Bag 34, Hobart TAS 7000

[5]Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China, 210044

[6]State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology, Harbin 150090

*Corresponding author phone: 86-411 8472-3185; fax: 86-411 8472-3185

E-mail address: degaowang@dlmu.edu.cn (De-Gao Wang)

Keywords: wastewater analysis; tobacco use; Monte Carlo simulation; trans-3'-hydroxycotinine; cotinine

Word count: 2900

Abstract

Objective: A new method of Monte Carlo simulation was developed to assess the uncertainty and variability of tobacco consumption through wastewater analysis in a city.

Methods: A total of 11 wastewater treatment plants (WWTPs) served 2.2 million population representing 83% of urban population in Dalian were selected and sampled. By detection and quantification of principal metabolites of cotinine (COT) and trans-3'-hydroxycotinine (OH-COT) in raw wastewater, back calculation of tobacco use in the population of WWTPs can be realized.

Results: COT and OH-COT were detected in the entire set of analysed samples with an average concentration of 2.33±0.30 and 2.76±0.91 μg/L, respectively. The mass load of absorbed NIC during the sampling period ranged from 0.25 to 4.22 mg/day/capita with an average of 1.92 mg/day/capita. The average consumed cigarette number is 14.6 cigarettes/day/capita for active smoker. Uncertainty and variability analysis by Monte Carlo simulation showed that the numbers of cigarettes per smokers in Dalian varied between 10 and 27 per day. The value of obtained from sewage analysis shows good agreement with the number of cigarettes investigated by epidemiological research.

Conclusions: Sewage-based epidemiology may be a useful additional tool for the large-scale monitoring of patterns of nicotine use. For the first time, probabilistic method used to assess the uncertainty can strength the reliability of estimation on the base of sewage epidemiology.

Strengths and limitations of this study

This paper describes a probabilistic method to assess the uncertainty of tobacco use based on wastewater analysis.

The approach offers a model to estimate of tobacco consumption in a city.

The main limitation to this study is that sewage-based epidemiology gives no information on demographic characteristics of smokers.

Introduction

China is the largest consumer of tobacco in the world, with an estimated 301 million smokers in 2010 [1]. The average cigarette consumption per smoker has remained relatively consistent during this century (estimated at 14.8 cigarettes per day in 2002 and 14.2 cigarettes per day in 2010) [2 3]. In May 2015, the Chinese government approved a increased in the tax applied to tobacco products, however, more tobacco-control programs and initiatives are needed to reduce the smoking-related disease burden in China [4]. Monitoring tobacco consumption is essential for evaluating the effectiveness of tobacco-control programs and initiatives.

To investigate population smoking prevalence, cross-sectional household surveys are the main sources of smoking prevalence estimates. The reliability of these estimates relies on large representative samples of the population providing accurate information on their tobacco use. In addition, the conventional method of questionnaire and socio-epidemiological surveys are labor-intensive and time-consuming [5]. A new method of sewage epidemiology based on wastewater analysis has been developed to investigate tobacco consumption. The approach is based on the principle that the metabolites of nicotine (NIC) ingested in human body are excreted with urine into urban sewer networks. By detection and quantification of principal metabolites of cotinine (COT) and trans-3'-hydroxycotinine (OH-COT) in raw wastewater, back calculation of tobacco use in population of a wastewater treatment plant (WWTP) catchment area can be realized. The wastewater analysis of active NIC excretion products has been recently used to estimate tobacco consumption in several cities in Spain [6] and Italy [7].The results show good agreement with prevalence data from national epidemiological surveys and demonstrate

the potential of the approach to complement existing socio-epidemiological methods [8]. The main advantage of sewage epidemiology provides evidence-based, objective, and real-time estimates of drug consumption in a defined population [9-17]. However, this approach is also subjective to a number of uncertainties associated with the different steps involved [18-20]. The key uncertainties are those related to the sampling of wastewater [21 22], the chemical analysis [23], the method to back-calculated tobacco use, and the estimation of the size of the population [24-26], human urinary excretion of metabolites [16 27 28]. Gaussian error propagation has become widely used for to calculate the uncertainty [19 20 29]. However, this approach analyzes the uncertainty of average values of estimation while not the variability of estimation. There are few studies to assess the uncertainty and variability of sewage epidemiology approach using the probabilistic method. As a representative of probabilistic approach, Monte Carlo simulation can quantify model inputs in estimation model, and therefore assess the variability and uncertainty of tobacco consumption. Hence, the major objectives of the present study are to integrally estimate the uncertainties of tobacco consumption through wastewater analysis using the sampling and analytical data collected from 11 WWTPs in Dalian, China.

Methods

Wastewater sampling and analysis

WWTPs were selected for sampling to achieve a wide geographical distribution in Dalian, China. In total, 11 plants were included, with all having more than 2.2 million people connected to these WWTPs. Twenty-four hour composite samples of raw wastewater were collected in polyethylene terephthalate containers by the staff of each

plant in June 2015. A standard questionnaire was developed to systematically gather relevant information for each WWTP including the population size, special events, flow variations, sampling mode, and frequency. A solid phase extraction with reverse-phase cartridges was applied for the clean-up of the wastewater sample. The liquid chromatography coupled with tandem mass-spectrometry was used to analyze the samples. The details of sample treatment, analysis, and chemicals used in this study can be found in the supporting information.

Back-calculation of NIC and tobacco consumption

For smoked NIC, the excretion percentage of COT and OH-COT, the molecular weight ratio between NIC and the metabolite, the flow rate and population for one WWTP, were used to calculate the mass of NIC consumed per day (m_{NIC}), as summarized in Equation 1:

$$m_{NIC} = \frac{F \times M_{NIC}}{P \times (x_{COT} + x_{OH\text{-}COT})} \left(\frac{C_{COT}}{M_{COT}} + \frac{C_{OH\text{-}COT}}{M_{OH\text{-}COT}}\right) \quad 1$$

where C_{COT} and $C_{OH\text{-}COT}$ are the concentrations of COT and OH-COT, F is the flow rate of raw wastewater in WWTPs, M_{NIC}, M_{COT} and $M_{OH\text{-}COT}$ are molecular weights of NIC, COT and OH-COT, x_{COT} and $x_{OH\text{-}COT}$ are excretion percentage of COT and OH-COT relative to NIC (%).

In this study 11 WWTPs were selected in Dalian, the average of NIC consumed (m_{NIC_Mean}) should be calculated by the each mass load ($m_{NIC,i}$) multiplying their weight (W_i) (Equation 2).

$$m_{NIC\_Mean} = \sum_{i=1}^{n} m_{NIC,i} W_i \quad 2$$

The weight for each WWTP can be calculated based between their population served and total population for all WWTPs (Equation 3).

$$W_i = \frac{P_i}{\sum_{i=1}^{n} P_i} \quad (3)$$

From an epidemiological and public health perspective, a desirable goal would be to estimate the number of cigarettes consumed per smoker rather than the net mass of NIC within a community. Recognizing this, the consumed cigarettes number per smoker can be calculated based on Equation 4:

$$n = \frac{m_{\text{NIC\_Mean}}}{D \times Y \times R_{15} \times x_{\text{Smoker}}} \quad (4)$$

where n is the number of cigarettes smoked per smoker, D is the content of NIC in an average cigarette, Y is the average yield of NIC uptake during smoking, R_{15} is the ratio of population aged 15 and older (15+) in Dalian, and x_{Smoker} is the smoker ratio (15+) in China. The average value of D is 0.90±0.15 mg of NIC per cigarette averaged with the most popular brands of cigarettes sold in Chinese market. The value of Y was determined to be 0.66±0.19 between the amount measured for that cigarette brand by standard Federal Trade Commission method and metabolites of NIC in urine [30]. The values of 91% and 24% for R_{15} and x_{Smoker} were used to calculate the smoking cigarette number.

Urinary excretion profile of NIC

Studies of NIC metabolism have been conducted over several decades. Many studies reported excretory profiles of NIC administered through smoking, oral ingestion, and transdermal administration [31-39]. The analysis of compiled data suggests that the urinary

excretion profile is quite similar when NIC is either administered through smoking or transdermal patches; however, the oral ingestion has higher extraction rates of OH-COT and COT (Table S2). The main urinary NIC metabolites are NIC, COT, and OH-COT and their conjugated forms [40]. The conjugated forms of these metabolites are completely transformed to the free form by β-glucuronidase enzymes from faecal bacteria in the raw wastewater [7]. Therefore, total amount (% of free and conjugated forms) of OH-COT and COT accounts for 43.4±13.8% and 32.3±8.0% of the total amount of NIC equivalents excreted when smoking is the main route for tobacco use. These mean excretion percentages are used to calculate mass of NIC consumed.

Uncertainty Analysis

The calculations of tobacco consumption are based on the average values of the input parameters. However, their implementations are complicated since the parameters cannot be treated as fixed-point values. Each parameter may take on a range of values depending on if the actual values are uncertain (e.g. concentrations of COT and OH-COT) or vary from person to person (e.g. excretion percentage of COT and OH-COT). The most effective quantification method for uncertainty and variability is to assign to each parameter a probability density function. Monte Carlo approach (Oracle Crystal Ball software, Version 7.3.1) allows the possibility of describing the uncertainty and variability associated with the input parameters and incorporating them into the estimates of tobacco consumption.

Selecting appropriate input distributions for uncertain and variable inputs is crucial to integrity of Monte Carlo simulation results. These distributions of density functions can be taken several different forms such as normal, lognormal, uniform, triangular. Normal

distribution can be used to represent a quantity for which the underlying mechanism can be described by the central limit theorem (CLT), such as the resultant of a large number of additive independent errors. In this study, we assumed the flow rates, population, the content of NIC in cigarettes, and the yield of NIC uptake during smoking are normal distributions. The flow rate was monitored by the online signal of the flow sensor in the inlet of the WWTP. The population served by the WWTPs is provided by the statistical data from the government, which manages and controls migration using the household register system and floating population registration system. Then, we assumed the normal distribution of flow rates and population with a conservative estimate of 10% relative standard deviation. A CLT can also be used as the basis for selecting a lognormal distribution to represent a quantity. The lognormal distribution has often been found to be a good representation of no-negative and results from the production of many random variation of the dilution of pollutant concentrations. Therefore, the concentrations and excretion percentage of COT and OH-COT can be taken as lognormal distribution. Table S3 and S4 show description of the Monte Carlo parameter distribution for tobacco use evaluation through smoking.

Sensitivity analysis shows how much each input parameter contributes to the uncertainty or variability of the estimations. In turn, it adds information on which portion of the variability is from natural fluctuation versus how much is caused by lack of knowledge. The contribution of an assumption to the total variance of the forecast is obtained as the square of the correlation coefficient normalized to the sum of the squared correlation coefficients adjusted to 100%.

Results and discussion

Concentrations in wastewater

NIC, COT, and OH-COT were detected in the entire set of analysed samples (frequency of detection 100%) with average concentrations of 11.5±8.73, 2.33±0.30 and 2.76±0.91 µg/L (see Table 1), respectively. As expected, OH-COT concentrations are higher than those of COT, in concordance with the known metabolic fate of NIC in humans. The values are highly correlated (R^2 = 0.90), with that of OH-COT being ~1.2 times greater than those of COT, in agreement with the NIC metabolism data (~1.3 times). The concentrations measured in the samples collected on 11 WWTPs are very similar and stable as expected. In contrast, parent NIC concentrations are higher than expected considering only urinary excretion and show random variability during different WWTPs. Cigarette butts and direct disposal of ash contribute to the other sources of NIC in wastewater, which indicates the human metabolites of COT and OH-COT are best biomarkers to estimate tobacco consumption. Actually, NIC concentrations do not correlated well with those of COT (R^2 = 0.69) and OH-COT (R^2 = 0.72).

These concentrations are in the same range of those measured by Lopes et al. in Lisbon, Portugal (1.1–3.5 µg/L for COT)[41] and slightly higher than those reported by Rodríguez-Álvarez et al. in Spain (2.2–5.9 µg/L, 1.0–1.6µg/L, and 1.8–2.8 µg/L for NIC, COT, and OH-COT)[6]. The concentrations can be influenced by many factors, such as tobacco prevalence, flows arriving to the WWTPs.

Table 1. The concentration of NIC, COT, and OH-COT in wastewater, flow rate, and served population, mass loads, and smoking cigarette number for 11 WWTPs in Dalian, China

WWTPs	Population ($\times 10^4$)	Flow rate ($\times 10^4$ m^3/d)	$C_{OH\text{-}COT}$ (μg/L)	C_{COT} (μg/L)	C_{NIC} (μg/L)	Mass load (mg/day/capita)	Cigarette number (cigarettes/day/smoker)
MYI	35	10	2.97±0.32	2.52±0.32	15.9±5.47	1.82	13.8
MER	25	8.15	2.99±0.33	2.55±0.32	6.81±5.84	2.10	15.9
LSH	22	5.5	2.04±0.21	1.76±0.16	4.77±1.21	1.10	8.4
FJZ	10	0.7	1.57±0.79	1.47±0.63	10.2±9.25	0.25	1.9
LHT	27	8	2.30±0.14	1.79±0.07	3.88±1.10	1.40	10.7
QSH	19.5	3.4	3.50±0.10	3.56±0.65	18.6±8.29	1.44	10.9
SEG	20	8	3.51±0.05	3.06±0.23	13.5±1.99	3.06	23.2
CYI	18.8	8	2.09±0.66	1.35±0.80	7.77±7.15	1.69	12.9
CER	28.2	12	3.22±0.30	2.97±0.06	14.7±0.62	3.07	23.3
HTX	8.1	2.5	1.77±0.28	1.03±0.22	0.67±0.74	1.00	7.6
XJH	6.5	3	4.36±0.83	3.51±0.62	29.2±3.17	4.22	32.0

Mass load of NIC consumed and uncertain analysis

The mass load of absorbed NIC during the sampling period ranges from 0.25 mg/day/capita to 4.22 mg/day/capita with an average of 1.92 mg/day/capita (Table 1). The mass load of NIC consumed reported in this work are similar than those reported in Spain (1.8 mg/day/capita) and lower than in Italy (3.42 mg/day/capita). These differences in the mass load of NIC consumed may be explained by the tobacco consumed behavior or habits.

Monte Carlo simulation shows the probability distribution density of the NIC mass load in Dalian (Figure 1). The results show the 10^{th} percentile, the central tendency of consumption (50^{th} percentile) and the reasonable maximum mass load (90^{th} percentile). The value of m_{NIC_Mean} was estimated to be 2.04 mg/day/capita, ranged from 1.56 mg/day/capita to 2.39 mg/day/capita in Dalian. The results reveal that the uncertainty of mass load, as defined by the ratio of the percentile 90^{th} to the percentile 10^{th}, is 1.53.

Figure 2 shows the sensitivity analysis for m_{NIC_Mean} and the eight most significant variables influencing mass load (>1%) are presented. The most two significant parameters influencing mass load are x_{COT} and $x_{OH\text{-}COT}$, which contribute 52% and 31% of total uncertainty of mass load, respectively. Totally, the population, flow rate, and concentration only contribute 17% of variance. In this analysis, it can be seen that variability of extraction percentage are far more important than the uncertainty about population, flow rate, and concentration to overall variance in mass load. The variability of extraction percentage for smokers refers to inter-individual heterogeneity with respect to different extraction rates of OH-COT and COT. The mechanisms underlying these

differences of urinary of NIC metabolites likely involve both genetic and behavioral factors [42]. Variability is an inherent property of a system under study and therefore is irreducible. While the uncertainty of population, flow rate, and concentration due to random or systematic error in measurement can be reduced with better data and/or better models. The main uncertainty of estimation results originates from the inter-individual variability, which cannot be reduced through improving the measurement technique.

Consumption of tobacco and uncertain analysis

Given that mean value of 0.90 mg/cigarette of NIC content in Chinese market and 66% yield of NIC uptake, the average value of 3.0 cigarettes/day/capita for in Dalian are obtained. According to the active smoker accounting for 24% of population (15+) in China and 91% of adults (15+) in Dalian, the consumed cigarettes number is 14.6 cigarettes/day/capita for active smoker. Global Adult Tobacco Survey report show that manufactured cigarette smokers (15+) consumed the average 14.2 manufactured cigarettes per day in 2010 in China [3]. The smoking prevalence investigation among urban residents of Liaoning province including Dalian showed that average amount cigarettes smoking were 14.2±0.3 per day in 2009 [43]. Hence, the value of obtained from sewage analysis shows good agreement between the numbers of cigarettes smoked as estimated from epidemiological survey data.

The probability distribution density for consumed number of cigarettes per smoker in Dalian is shown in Figure 3. Since estimation is calculated as the product of several probability distributions, the mass load distribution shows in Figure 3 to be approximately lognormal as result of the CLT. The distribution for consumption indicates that there are approximately 10%, 50%, and 90% chances lower than 10, 16, and 27 per

day in Dalian. It can be seen that the median (50% percentile) of number of cigarettes is 16, indicating the smokers in Dalian smoke 0.8 pack of cigarettes per day. The results reveal that the uncertainty and variability of number of cigarettes estimated is 2.0. It can be concluded that the smokers in Dalian smoke from 0.55 to 1.1 packs of cigarettes per day. The probability distribution of smoking cigarette number is consistent with the questionnaire review from Jiangsu Province in 2010, which showed the median number of cigarettes consumed per day was 15 and ranged from 10 to 20 for the 25th and 75th percentile [44].

The sensitivity analysis for number of cigarettes smoked shows that the four parameters of Y, D, $x_{OH\text{-}COT}$, and x_{COT} contributed to 63%, 20%, 9%, and 5% of variance, respectively (Figure 4). The contribution of the each other parameter is lower than 1%. The content of NIC in one cigarette is the most important parameters than other factors, which almost contribute half of uncertainty of number of smoking cigarettes. Due to inter-individual heterogeneity of three parameters of Y, $x_{OH\text{-}COT}$, and x_{COT}, further study on improving the estimating the true values of D should be done to reduce the uncertainty of estimation for cigarette consumption.

In summary, the present study provides a systematic assessment of tobacco use in a city based on sewage epidemiology. Generally, the number of cigarettes consumed per capita agrees quite well with those derived from a separate epidemiology survey. Monte Carlo simulation shows that the uncertainty of mass load and cigarette number are 1.5 and 2.0. The results of sensitivity analysis show that the inter-individual variability plays an important role on uncertainty of estimation results. The data in the present study are the first published estimates generated using this tool for tobacco consumption in a

Chinese city. Although sewage epidemiology is a generally objective and real-time technique, there are also limitations. This approach gives no information on demographic characteristics of smokers. Further work is required to develop new techniques to reduce the uncertainty of consumption. Sewage epidemiology is a potential approach to estimate consumption of nicotine, and could complement conventional socio-epidemiological surveys, particularly when it comes to the real-time monitoring of the impact of new tobacco control policies.

Acknowledgements

Thanks for the discussion with Dr. Lian-Zheng Yu from the Department of Noncommunicable Chronic Disease Prevention, Liaoning Center for Disease Control and Prevention.

Contributors

De-Gao Wang designed whole study, analysed data and wrote and revised the paper. Qian-Qian Dong pretreated the samples. Juan Du pretreated the samples. Shuo Yang collected the samples. Yun-Jie Zhang designed the model. Guang-Shui Na analysed the samples. Stuart G. Ferguson analysed the data and revised the paper. Zhuang Wang revised the draft paper. Tong Zheng revised the draft paper.

Funding

This study was supported by the Fundamental Research Funds for the Central Universities, Open Fund by Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (KHK1309), A Project Funded by the Priority

Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and Open Project of State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology (QAK201503).

Provenance and peer review

Not commissioned; externally peer reviewed.

Competing interests

None.

Data sharing statement

No additional data are available.

References

1. Li Q, Hsia J, Yang G. Prevalence of Smoking in China in 2010. New England Journal of Medicine 2011;**364**(25):2469-70.
2. Sung HY, Wang L, Jin S, et al. Economic burden of smoking in China, 2000. Tobacco Control 2006;**15**(3):i5-i11.
3. WHO. Global Adult Tobacco Survey (GATS) China 2010 Country Report.: Chinese Center for Disease Control and Prevention, 2011.
4. Wang JB, Jiang Y, Wei WQ, et al. Estimation of cancer incidence and mortality attributable to smoking in China. Cancer Causes and Control 2010;**21**(6):959-65.
5. Gartner C. Flushing out smoking: measuring population tobacco use via wastewater analysis. Tobacco Control 2015;**24**(1):1-2.

6. Rodríguez-Álvarez T, Rodil R, Rico M, et al. Assessment of local tobacco consumption by liquid chromatography–tandem mass spectrometry sewage analysis of nicotine and its metabolites, cotinine and trans-3′-hydroxycotinine, after enzymatic deconjugation. Analytical Chemistry 2014;**86**(20):10274-81.

7. Castiglioni S, Senta I, Borsotti A, et al. A novel approach for monitoring tobacco use in local communities by wastewater analysis. Tobacco Control 2015;**24**(1):38-42.

8. Reid MJ, Langford KH, Grung M, et al. Estimation of cocaine consumption in the community: a critical comparison of the results from three complimentary techniques. Bmj Open 2012;**2**(6):1-9.

9. Li J, Hou L, Du P, et al. Estimation of amphetamine and methamphetamine uses in Beijing through sewage-based analysis. The Science of the total environment 2014;**490**:724-32.

10. Du P, Li K, Li J, et al. Methamphetamine and ketamine use in major Chinese cities, a nationwide reconnaissance through sewage-based epidemiology. Water Research 2015;**84**:76-84.

11. Castiglioni S, Bijlsma L, Covaci A, et al. Evaluation of uncertainties associated with the determination of community drug use through the measurement of sewage drug biomarkers. Environmental Science and Technology 2012;**47**(3):1452-60.

12. Reid MJ, Langford KH, Mørland J, et al. Quantitative assessment of time dependent drug-use trends by the analysis of drugs and related metabolites in raw sewage. Drug and Alcohol Dependence 2011;**119**(3):179-86.

13. van Nuijs ALN, Castiglioni S, Tarcomnicu I, et al. Illicit drug consumption estimations derived from wastewater analysis: A critical review. Science of The Total Environment 2011;**409**:3564-77.

14. van Nuijs ALN, Mougel J-F, Tarcomnicu I, et al. Sewage epidemiology - A real-time approach to estimate the consumption of illicit drugs in Brussels, Belgium. Environment International 2011;**37**(3):612-21.

15. Postigo C, Lopez de Alda MJ, Barceló D. Analysis of drugs of abuse and their human metabolites in water by LC-MS^2: A non-intrusive tool for drug abuse estimation at the community level. TrAC Trends in Analytical Chemistry 2008;**27**(11):1053-69.

16. Ort C, van Nuijs AL, Berset JD, et al. Spatial differences and temporal changes in illicit drug use in Europe quantified by wastewater analysis. Addiction 2014;**109**(8):1338-52.

17. Thomas KV, Bijlsma L, Castiglioni S, et al. Comparing illicit drug use in 19 European cities through sewage analysis. Science of The Total Environment 2012;**432**:432-9.

18. Lai FY, Ort C, Gartner C, et al. Refining the estimation of illicit drug consumptions from wastewater analysis: Co-analysis of prescription pharmaceuticals and uncertainty assessment. Water Research 2011;**45**(15):4437-48.

19. Mathieu C, Rieckermann J, Berset J-D, et al. Assessment of total uncertainty in cocaine and benzoylecgonine wastewater load measurements. Water Research 2011;**45**(20):6650-60.

20. Castiglioni S, Bijlsma L, Covaci A, et al. Evaluation of uncertainties associated with the determination of community drug use through the measurement of sewage drug biomarkers. Environmental Science and Technology 2013;**47**(3):1452-60.

21. Ort C, Eppler JM, Scheidegger A, et al. Challenges of surveying wastewater drug loads of small populations and generalizable aspects on optimizing monitoring design. Addiction 2014;**109**(3):472-81.

22. Ort C, Lawrence MG, Rieckermann J, et al. Sampling for pharmaceuticals and personal care products (PPCPs) and illicit drugs in wastewater systems: are your conclusions valid? A critical review. Environmental Science and Technology 2010;**44**(16):6024-35.

23. Thai PK, Jiang G, Gernjak W, et al. Effects of sewer conditions on the degradation of selected illicit drug residues in wastewater. Water Research 2014;**48**:538-47.

24. Chen C, Kostakis C, Gerber JP, et al. Towards finding a population biomarker for wastewater epidemiology studies. Science of the Total Environment 2014;**487**:621-28.

25. Daughton CG. Real-time estimation of small-area populations with human biomarkers in sewage. Science Of the Total Environment 2012;**414**:6-21.

26. Lai FY, Anuj S, Bruno R, et al. Systematic and day-to-day effects of chemical-derived population estimates on wastewater-based drug epidemiology. Environmental Science and Technology 2014;**49**(2):999-1008.

27. Khan U, Nicell JA. Refined sewer epidemiology mass balances and their application to heroin, cocaine and ecstasy. Environ Int 2011;**37**(7):1236-52.

28. Khan U, Nicell JA. Sewer epidemiology mass balances for assessing the illicit use of methamphetamine, amphetamine and tetrahydrocannabinol. Science of the Total Environment 2012;**421**:144-62.

29. Brewer AJ, Ort C, Banta-Green CJ, et al. Normalized diurnal and between-day trends in illicit and legal drug loads that account for changes in population. Environmental Science and Technology 2012;**46**(15):8305-14.

30. St.Charles FK, Kabbani AA, Borgerding MF. Estimating tar and nicotine exposure: Human smoking versus machine generated smoke yields. Regulatory Toxicology and Pharmacology 2010;**56**(1):100-10.

31. Byrd GD, Chang KM, Greene JM, et al. Evidence for urinary excretion of glucuronide conjugates of nicotine, cotinine, and trans-3'-hydroxycotinine in smokers. Drug Metabolism and Disposition 1992;**20**(2):192-7.

32. Gunilla A, Gunilla B, Margareta C. Oral mucosal changes and nicotine disposition in users of Swedish smokeless tobacco products: a comparative study. Journal of Oral Pathology and Medicine 1994;**23**(4):161-67.

33. Byrd GD, Robinson JH, Caldwell WS, et al. Comparison of measured and FTC-predicted nicotine uptake in smokers. Psychopharmacology 1995;**122**(2):95-103.

34. Andersson G, Vala EK, Curvall M. The influence of cigarette consumption and smoking machine yields of tar and nicotine on the nicotine uptake and oral mucosal lesions in smokers. Journal of Oral Pathology and Medicine 1997;**26**(3):117-23.

35. Byrd GD, Davis RA, Caldwell WS, et al. A further study of FTC yield and nicotine absorption in smokers. Psychopharmacology 1998;**139**(4):291-99.

36. Meger M, Meger-Kossien I, Schuler-Metz A, et al. Simultaneous determination of nicotine and eight nicotine metabolites in urine of smokers using liquid chromatography–tandem mass spectrometry. Journal of Chromatography B 2002;**778**(1-2):251–61.

37. Benowitz NL, Pomerleau OF, Pomerleau CS, et al. Nicotine metabolite ratio as a predictor of cigarette consumption. Nicotine and Tobacco Research 2003;**5**(5):621-24.

38. Mcguffey JE, Wei B, Bernert JT, et al. Validation of a LC-MS/MS method for quantifying urinary nicotine, six nicotine metabolites and the minor tobacco alkaloids--anatabine and anabasine--in smokers' urine. Plos One 2014;**9**(9).

39. Benowitz NL, Fong I, Gupta S. Nicotine metabolic profile in man: comparison of cigarette smoking and transdermal nicotine. Journal of Pharmacology and Experimental Therapeutics 1994;**268**(1):296-303.

40. Jacobson GA, Ferguson SG. Relationship between cotinine and trans-3′-hydroxycotinine glucuronidation and the nicotine metabolite ratio in Caucasian smokers. Biomarkers 2014;**19**(8):679-83.

41. Lopes A, Silva N, Bronze MR, et al. Analysis of cocaine and nicotine metabolites in wastewater by liquid chromatography-tandem mass spectrometry. Cross abuse index patterns on a major community. Science of the Total Environment 2014;**487**:673-80.

42. Derby K, Cuthrell K, C, Carmella S, et al. Nicotine metabolism in three ethnic/racial groups with different risks of lung cancer. Cancer Epidemiology, Biomarkers and Prevention 2008;**17**(12):3526-35.

43. Yu L, Mei D, An X, et al. Analysis on diffierence in smoking prevalence among urban residents in Liaoning Province, 1999-2009. Chinese Journal of Disease Control and Prevention 2012;**16**(10):842-45.

44. Su J, QIn Y, Xiang Q, et al. Smoking behaviors and tobacco consumption of smokers among residents in Jiansu Province. Chinese Journal of Public Health 2014;**30**(11):1368-72.

Figure captions

Figure 1. Distribution of estimated mass load of NIC (mg/day/capita) in Dalian: mean, median, 10th, and 90th percentiles.

Figure 2. The results of sensitivity analysis for NIC mass load. x_{COT} and $x_{OH\text{-}COT}$ are excretion percentage of COT and OH-COT relative to NIC (%); P_{cer} and F_{cer} are the population and flow rate in WWTP of CER; C_{COT_cyi} is the concentration of COT in WWTP of CYI; P_{seg} and F_{seg} are the population and flow rate in WWTP of SEG; P_{myi} is the population in WWTP of MYI.

Figure 3. Distribution of estimated number of cigarette consumed (cigarettes/day/smoker) in Dalian: mean, median, 10th and 90th percentiles.

Figure 4. The results of sensitivity analysis for number of cigarettes consumed. x_{COT} and $x_{OH\text{-}COT}$ are excretion percentage of COT and OH-COT relative to NIC (%); Y is the yield of NIC during smoking, and D is the content of NIC in one cigarette.

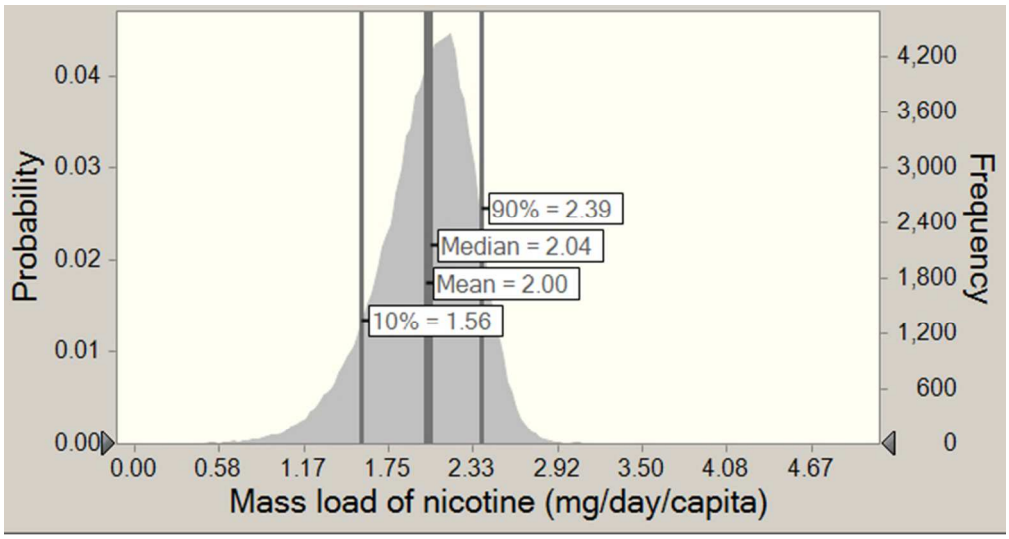

Figure 1. Distribution of estimated mass load of NIC (mg/day/capita) in Dalian: mean, median, 10th, and 90th percentiles.
56x29mm (300 x 300 DPI)

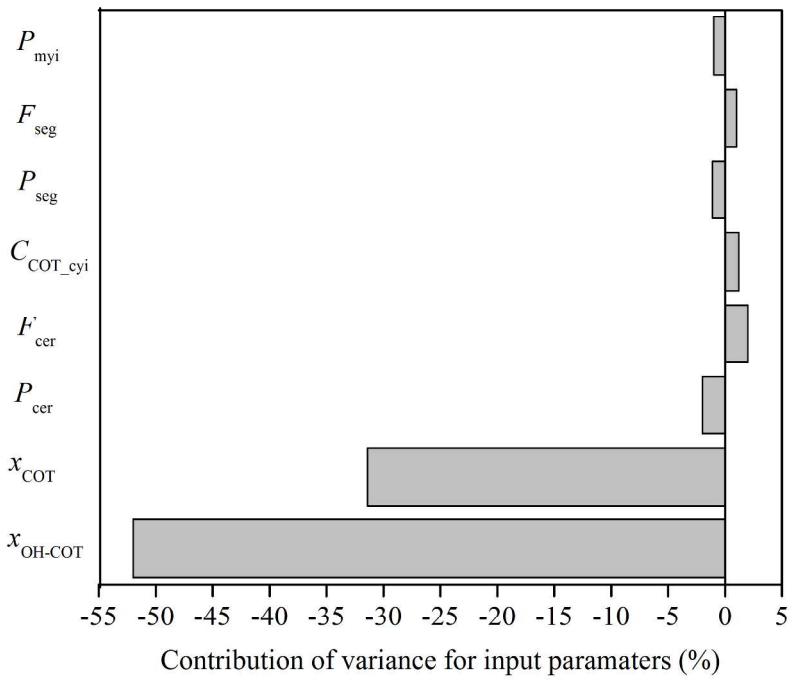

Figure 2. The results of sensitivity analysis for NIC mass load. xCOT and xOH-COT are excretion percentage of COT and OH-COT relative to NIC (%); Pcer and Fcer are the population and flow rate in WWTP of CER; CCOT_cyi is the concentration of COT in WWTP of CYI; Pseg and Fseg are the population and flow rate in WWTP of SEG; Pmyi is the population in WWTP of MYI.
254x203mm (300 x 300 DPI)

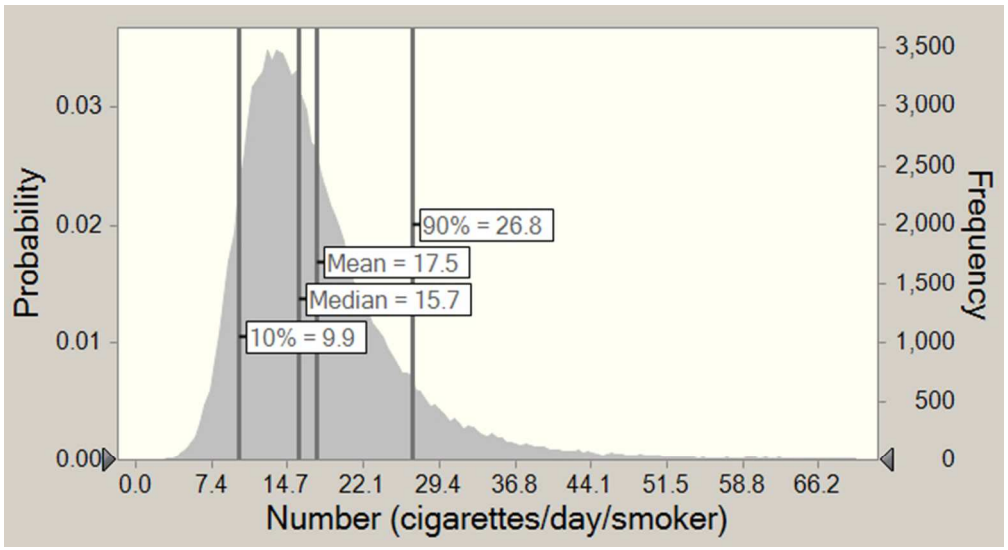

Figure 3. Distribution of estimated number of cigarette consumed (cigarettes/day/smoker) in Dalian: mean, median, 10th and 90th percentiles.
56x30mm (300 x 300 DPI)

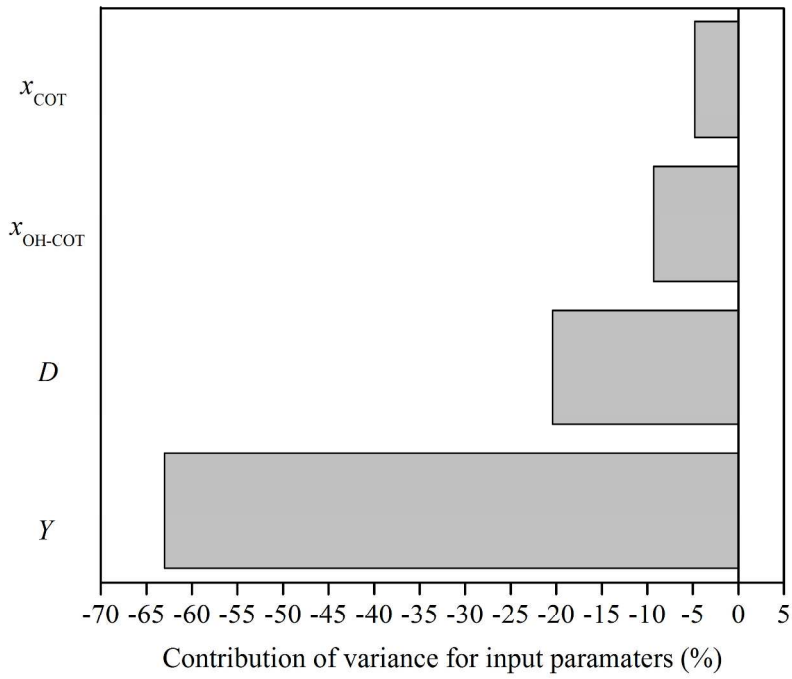

Figure 4. The results of sensitivity analysis for number of cigarettes consumed. xCOT and xOH-COT are excretion percentage of COT and OH-COT relative to NIC (%);Y is the yield of NIC during smoking, and D is the content of NIC in one cigarette.
254x203mm (300 x 300 DPI)

Supporting Information

Monte Carlo simulation assess variability and uncertainty of tobacco consumption in a city by sewage epidemiology

De-Gao Wang[1†*], Qian-Qian Dong[1], Juan Du[1], Shuo Yang[1], Yun-Jie Zhang[2], Guang-Shui Na[3], Stuart G. Ferguson[4], Zhuang Wang[5], Tong Zheng[6]

[1]School of Environmental Science and Technology, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[2]Department of Mathematics, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[3]National Marine Environmental Monitoring Center, Dalian, Liaoning, China, 116023

[4]Faculty of Health, School of Medicine, University of Tasmania Private Bag 34, Hobart TAS 7000

[5]Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China

[6]State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology, Harbin 150090

*Corresponding author phone: 86-411 8472-3185; fax: 86-411 8472-3185

E-mail address: degaowang@dlmu.edu.cn (De-Gao Wang)

Chemicals

NIC and cotinine-d3 (COT-d3) were purchased from Sigma−Aldrich Inc. (Saint Louis, MO, U.S.A). Nicotine-d3 (NIC-d3) was purchased from Toronto Research Chemicals Inc. (TRC, ON, Canada). COT was purchased from ChromaDex, Inc. (ChromaDex, CA, USA). OH-COT and trans-3′-hydroxycotinine-d3 (OH-COT-d3) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Individual stock solutions were prepared in acetonitrile and diluted as necessary. HPLC-grade methanol and acetonitrile (ACN) were purchased from J&K Chemical (J&K, China). Ultrapure water was obtained from a Milli-Q water generator (Millipore, USA).

Analysis

Raw wastewater (20 mL) was passed through a glass filter (0.45 μm) to removal particles. Then, deuterated labeled internal standards (20 ng) for quantification were added and a solid phase extraction (SPE) clean-up set with reverse-phase cartridges (Cleanert PEP-2, 1 mL, 60 mg, China) was applied. The cartridges were conditioned with consecutively methanol (2 mL) and Milli-Q water (2 mL). The extracts were allowed to enter the preconditioned SPE cartridge connected to a vacuum pump. The cartridge was washed with 1 mL of Milli-Q water. After drying of the cartridges, the elution was performed with 1 mL of methanol. The eluate was evaporated and redissolved in 1 mL of ACN for final analysis.

The liquid chromatography coupled with tandem mass-spectrometry (DSQ600, Canada) (LC–MS/MS) was used to analyze the samples. The injection volume was 10 μL. Targeted analytes were chromatographically separated using a Waters Symmetry Shield

C18 column (3.5 μm, 150 mm×2.1 mm, Waters, USA) at a flow rate of 0.4 mL/min. Eluents used for the separation included 5 mmol/L NH_4AC in H_2O (eluent A) and 5 mmol/L NH_4AC in methanol (eluent B). The gradient used as follows: Start as 0 min with 20% B and 80% A, held for 2 min; 2-6 min linear rate to 100% B, held for 2 min; 8-10 min linear rate to 20% B, 10-11 min return to initial conditions; 20% B.

Scheduled multiple reaction monitoring with positive electrospray ionisation were operated to identify and quantify the masses of analytes. The collision energy and tube lens offset were optimised for each analyte and surrogate separately. Positive ionization electrospray was used with nitrogen as the cone and desolvation gas. Identification and quantification was performed using two characteristic transitions for the analysed compounds (Table S1).

Quantification was performed using a seven-point calibration curve spanning the range of anticipated analyte concentrations in the samples. The method precision and recovery were tested for both matrices, by spiking three parallel samples of Milli-Q water and wastewater with a mixture of all analytes at environmentally relevant concentrations. The recoveries of analytes varied between 64% and 93%. All concentrations of analytes were corrected by the recoveries. The limits of detection were estimated to be 0.020 μg/L, 0.012 μg/L, and 0.014 μg/L for NIC, COT, and OH-COT, respectively.

Table S1 Target compounds and stable isotope standards with the retention times, the mass spectrometric SRM ion transitions, de-clustering potential (DP) and collision energies (CE) for analysis, and the limits of detection (LOD) and recoveries for target analytes

Compound	Retention Time	SRM Transition	DP(V)	CE(V)	LOD (μg/L)	Recovery (%)
NIC	2.96	163→117	84	35	0.020	64
NIC-d3	2.96	166→106	48	22	-	-
COT	1.67	177→80	60	27	0.012	72
COT-d3	1.67	180→80	60	28	-	-
OH−COT	1.60	193→80	56	29	0.014	93
OH−COT-d3	1.60	196→80	56	48	-	-

Table S2 Distribution of urinary nicotine metabolites as percentage of the total amount of metabolites excreted (mean and standard deviation)

Reference	Route	Dose (mg)	n	Age (yr)	Duration (h)	OH-COT	OH-COT-G	COT	COT-G	NIC	NIC-G	OH-COT-T	COT-T	NIC-T
Byrd et al.(1992)	Smoked	7.2-48	11	29-47	24	35±8	9±4	13±4	17±7	10±4	3±2	45.3±10.4	29.8±7.8	12.7±3.6
Benowitz et al. (1994)	Smoked	34.2±21.4	12		24	39.1±12.5	7.8±5.9	13.3±3.1	15.8±7.8	10.4±4.4	4.6±2.4	46.9±18.4	29.1±10.9	15.0±7.3
Byrd et al.(1995)	Smoked	4.8-38.1	33	26-50	24	-	-	-	-	-	-	53.3±15.6	26.6±10.8	10.4±3.9
Andersson et al. (1997)	Smoked	25.0±10.7	91	38.0±12.1	24	36.1±10.6	22.8±10.0	9.2±2.6	14.0±5.4	9.4±5.7	4.5±2.5	58.9±20.6	23.2±8.0	13.9±8.2
Byrd et al.(1998)	Smoked	34±8	72	37±9	24	-	-	-	-	-	-	43.1±12.2	25.4±7.1	14.4±6.8
Meger et al.(2002)	Smoked	-	5	-	24	32.2±14	6.8±3.2	16.1±2.1	19.6±6.9	10.4±4.1	4.5±4	39.0±17.2	35.7±9.0	14.9±8.1
Benowitz et al. (2002)	Smoked	12.3±10.4	40	7.7±6.2	24	23.5±16.3	8.9±8.6	26.0±10.3	12.1±9.1	19.2±16.3	10.3±7.3	32.4±24.9	38.1±19.4	29.5±23.6
McGuffey et al. (2014)	Smoked	-	94	-	24	-	-	-	-	-	-	38.8±20.7	26.4±9.6	20.6±14.2
Average	Smoked	-	154	-	24	-	-	-	-	-	-	**43.4±13.8**	**32.3±8.0**	**19.4±8.7**
Benowitz et al. (1994)	Transdermal	21.5±4.5	12	-	24	37.0±10.8	7.9±4.7	14.9±4.6	15.4±7.9	11.1±4.3	5.3±3.3	44.9±15.5	30.3±12.5	16.4±7.6
Andersson et al. (1994)	Oral	9.3-124	54	43.3±10.7	24	41.6±10.6	19.4±11.0	7.9±2.2	8.9±4.6	8.3±5.7	3.0±1.8	61.0±21.6	16.8±6.8	11.3±7.5

Table S3. Monte Carlo parameter distribution for population, flow rate, and concentrations in 11 WWTP for estimation of tobacco use

WWTP	Symbol	Concentrations (μg/L)					Flow rate ($\times 10^4$ m^3/d)			Served population of WWTP ($\times 10^4$)			Weight		
		Type	Distribution	Symbol	Type	Distribution	Symbol	Type	Distribution	Symbol	Type	Distribution	Symbol	Type	Distribution
MYI	$C_{OH\text{-}COT\_myi}$	lognormal	2.97±0.32	C_{COT_myi}	lognormal	2.52±0.32	F_{myi}	normal	10±1.0	P_{myi}	normal	35±3.5	W_{myi}	point	0.22
MER	$C_{OH\text{-}COT\_mer}$	lognormal	2.99±0.33	C_{COT_mer}	lognormal	2.55±0.32	F_{mer}	normal	8.15±0.82	P_{mer}	normal	25±2.5	W_{mer}	point	0.18
LSH	$C_{OH\text{-}COT\_lsh}$	lognormal	2.04±0.21	C_{COT_lsh}	lognormal	1.76±0.16	F_{lsh}	normal	5.5±0.55	P_{lsh}	normal	22±2.2	W_{lsh}	point	0.08
FJZ	$C_{OH\text{-}COT\_fjz}$	lognormal	1.57±0.79	C_{COT_fjz}	lognormal	1.47±0.63	F_{fjz}	normal	0.7±0.07	P_{fjz}	normal	10±1.0	W_{fjz}	point	0.01
LHT	$C_{OH\text{-}COT\_lht}$	lognormal	2.30±0.14	C_{COT_lht}	lognormal	1.79±0.07	F_{lht}	normal	8±0.8	P_{lht}	normal	27±2.7	W_{lht}	point	0.13
QSH	$C_{OH\text{-}COT\_qsh}$	lognormal	3.50±0.10	C_{COT_qsh}	lognormal	3.56±0.65	F_{qsh}	normal	3.4±0.34	P_{qsh}	normal	19.5±1.95	W_{qsh}	point	0.10
SEG	$C_{OH\text{-}COT\_seg}$	lognormal	3.51±0.05	C_{COT_seg}	lognormal	3.06±0.23	F_{seg}	normal	8±0.8	P_{seg}	normal	20±2.0	W_{seg}	point	0.21
CYI	$C_{OH\text{-}COT\_cyi}$	lognormal	2.09±0.66	C_{COT_cyi}	lognormal	1.35±0.80	F_{cyi}	normal	8±0.8	P_{cyi}	normal	18.8±1.88	W_{cyi}	point	0.11
CER	$C_{OH\text{-}COT\_cer}$	lognormal	3.22±0.30	C_{COT_cer}	lognormal	2.97±0.06	F_{cer}	normal	12±1.2	P_{cer}	normal	28.2±2.82	W_{cer}	point	0.30
HTX	$C_{OH\text{-}COT\_htx}$	lognormal	1.77±0.28	C_{COT_htx}	lognormal	1.03±0.22	F_{htx}	normal	2.5±0.25	P_{htx}	normal	8.1±0.81	W_{htx}	point	0.03
XJH	$C_{OH\text{-}COT\_xjh}$	lognormal	4.36±0.83	C_{COT_xjh}	lognormal	3.51±0.62	F_{xjh}	normal	3±0.3	P_{xjh}	normal	6.5±0.65	W_{xjh}	point	0.09

Table S4. Monte Carlo parameter distribution of other input parameters for estimation of tobacco use

Parameter	Symbol	Units	Type	Distribution
Molecular weight of NIC	M_{NIC}	unitless	point	162
Molecular weight of COT	M_{COT}	unitless	point	176
Molecular weight of OH-COT	$M_{OH\text{-}COT}$	unitless	point	192
Excretion percentage of COT	x_{COT}	unitless	lognormal	32.3%±8.0%
Excretion percentage of OH-COT	$x_{OH\text{-}COT}$	unitless	lognormal	43.4%±13.8%
Population ratio in Dalian (>15)	R_{15}	unitless	point	0.91
Smoker ratio in China (>15)	X_{Smoker}	unitless	point	0.24
Content of NIC in one cigarette	D	mg	normal	0.90±0.15
Yield of NIC uptake	Y	unitless	normal	0.66±0.19

Research checklist

NO.

BMJ Open

Using Monte Carlo simulation to assess variability and uncertainty of tobacco consumption in a city by sewage epidemiology.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010583.R1
Article Type:	Research
Date Submitted by the Author:	06-Jan-2016
Complete List of Authors:	Wang, De-Gao; Dalian Maritime University, School of Environmental Science and Technology Dong, Qian-Qian; Dalian Maritime University, School of Environmental Science and Technology Du, Juan; Dalian Maritime University, School of Environmental Science and Technology Yang, Shuo ; Dalian Maritime University, School of Environmental Science and Technology Zhang, Yun-Jie ; Dalian Maritime University, Department of Mathematics Na, Guang-Shui ; National Marine Environmental Monitoring Center, Ferguson, Stuart ; University of Tasmania, Faculty of Health Science, School of Medicine Wang, Zhuang ; Nanjing University of Information Science and Technology, Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), School of Environmental Science and Engineering Zheng, Tong; Harbin Institute of Technology, State Key Laboratory of Urban Water Resource and Environment
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Addiction, Epidemiology, Smoking and tobacco
Keywords:	wastewater analysis, Monte Carlo simulation, cotinine, trans-3'-hydroxycotinine, tobacco use

Using Monte Carlo simulation to assess variability and uncertainty of tobacco consumption in a city by sewage epidemiology.

De-Gao Wang[1†*], Qian-Qian Dong[1], Juan Du[1], Shuo Yang[1], Yun-Jie Zhang[2], Guang-Shui Na[3], Stuart G. Ferguson[4], Zhuang Wang[5], Tong Zheng[6]

[1]School of Environmental Science and Technology, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[2]Department of Mathematics, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[3]National Marine Environmental Monitoring Center, Dalian, Liaoning, China, 116023

[4]Faculty of Health Science, School of Medicine, University of Tasmania Private Bag 34, Hobart TAS 7000

[5]Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China, 210044

[6]State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology, Harbin 150090

*Corresponding author phone: 86-411 8472-3185; fax: 86-411 8472-3185

E-mail address: degaowang@dlmu.edu.cn (De-Gao Wang)

Keywords: wastewater analysis; tobacco use; Monte Carlo simulation; trans-3'-hydroxycotinine; cotinine

Word count: 2900

Abstract

Objective: To use Monte Carlo simulation to assess the uncertainty and variability of tobacco consumption through wastewater analysis in a city.

Methods: A total of 11 wastewater treatment plants (WWTPs) (serving 2.2 million people; ~83% of urban population in Dalian) were selected and sampled. By detection and quantification of principal metabolites of nicotine, cotinine (COT) and trans-3'-hydroxycotinine (OH-COT), in raw wastewater, back calculation of tobacco use in the population of WWTPs can be realized.

Results: COT and OH-COT were detected in the entire set of samples with an average concentration of 2.33±0.30 and 2.76±0.91 μg/L, respectively. The mass load of absorbed NIC during the sampling period ranged from 0.25 to 4.22 mg/day/capita with an average of 1.92 mg/day/capita. Using these data, we estimated that smokers in the sampling area consumed an average of 14.6 cigarettes per day for active smoker. Uncertainty and variability analysis by Monte Carlo simulation was used to refine this estimate: the procedure concluded that smokers in Dalian smoked between 10 and 27 cigarettes per day. This estimate showed good agreement with estimates from epidemiological research.

Conclusions: Sewage-based epidemiology may be a useful additional tool for the large-scale monitoring of patterns of tobacco use. Probabilistic methods can be used to strengthen the reliability of estimated use generated from wastewater analysis.

Strengths and limitations of this study

This paper describes a probabilistic method to assess the uncertainty of tobacco use based on wastewater analysis.

The approach offers a model to estimate tobacco consumption in a city.

The main limitation to this study is that sewage-based epidemiology gives no information on demographic characteristics of smokers.

Introduction

China is the largest consumer of tobacco in the world, with an estimated 301 million smokers in 2010 [1]. The average cigarette consumption per smoker has remained relatively consistent during this century (estimated at 14.8 cigarettes per day [CPD] in 2002 and 14.2 CPD in 2010) [2 3]. In May 2015, the Chinese government approved an increased in the tax applied to tobacco products, however, more tobacco-control programs and initiatives are needed to reduce the smoking [4]. Monitoring tobacco consumption is essential for evaluating the effectiveness of tobacco-control programs and initiatives.

Cross-sectional household surveys are the main sources of cross-sectional smoking prevalence estimate. The reliability of these estimates relies on large representative samples of the population providing accurate information about their tobacco use. These methods are labor-intensive and time-consuming [5]. In part to overcome these issues, a new method of sewage epidemiology based on wastewater analysis has been developed to investigate tobacco consumption. The approach is based on the principle that the metabolites of nicotine (NIC) are excreted with urine into urban sewer networks. By the detection and quantification of the principal metabolites—cotinine (COT) and trans-3'-hydroxycotinine (OH-COT)—in raw wastewater, back calculation of tobacco use in the population of a wastewater treatment plant (WWTP) catchment area can be estimated. Such wastewater analysis procedures have been used to estimate tobacco consumption in several cities in Spain [6] and Italy [7].The results show good agreement with prevalence data from national epidemiological surveys and demonstrate the potential of the approach to complement existing socio-epidemiological methods [8].

The primary advantages of sewage epidemiology procedures are that they provide evidence-based, objective, and real-time estimates of drug consumption in a defined population [9-17]. However, this approach is also subjective to a number of uncertainties associated with the different steps involved [18-20]. The key uncertainties are those related to the sampling of wastewater [21 22], the chemical analysis [23], the method to back-calculated tobacco use, the estimation of the size of the population [24-26], and human urinary excretion of metabolites [16 27 28]. Gaussian error propagation has become widely used to calculate the uncertainty of estimations generated [19 20 29]. However, this approach analyzes the uncertainty of average values of estimation, not the variability of the estimation itself. As a representative of probabilistic approach, Monte Carlo simulations can quantify model inputs in estimation model, and therefore assess the variability and uncertainty of tobacco consumption. The objectives of the present study were to assess tobacco consumption using wastewater analyses in a region of China, and to use Monte Carlo simulations to estimate the uncertainties of the estimation of tobacco consumption generated.

Methods

Wastewater sampling and analysis

WWTPs were selected for sampling to achieve a wide geographical distribution in Dalian, China. In total, 11 plants were included, servicing a population of 2.2 million people. All 11 plants operated with the secondary wastewater treatment systems such as biological aerated filter, cyclic active sludge technology, sequencing batch reactor, and anoxic/oxic activated sludge process. Twenty-four hour composite samples of raw

wastewater were collected in polyethylene terephthalate containers by the staff of each plant on two consecutive weekdays in June 2015. A previous study suggested that the concentrations of COT and OH-COT were consistent between weekdays and weekends [7]. Each WWTP also provided information on special events that occurring during the sampling period, flow variations (as measured by a flow sensor in the inlet of the WWTP), sampling mode, and sampling frequency. The population served by each WWTP was estimated using statistical data from the government.

A solid phase extraction with reverse-phase cartridges was applied for the clean-up of the wastewater sample. Liquid chromatography coupled with tandem mass-spectrometry was then used to analyze the samples. The details of sample treatment, analysis (Table S1), and chemicals used in this study can be found in the supporting information.

Back-calculation of NIC and tobacco consumption

The percent of COT and OH-COT excreted, the molecular weight ratio between NIC and the metabolites, the flow rate, and the population served were used to calculate the mass load of NIC consumed in each WWTP per day (m_{NIC}), as summarized in Equation 1:

$$m_{NIC} = \frac{F \times M_{NIC}}{P \times (x_{COT} + x_{OH\text{-}COT})} \left(\frac{C_{COT}}{M_{COT}} + \frac{C_{OH\text{-}COT}}{M_{OH\text{-}COT}}\right) \quad (1)$$

where, C_{COT} and $C_{OH\text{-}COT}$ are the concentrations of COT and OH-COT, F is the flow rate of raw wastewater in WWTPs, M_{NIC}, M_{COT} and $M_{OH\text{-}COT}$ are molecular weights of NIC, COT, and OH-COT, x_{COT} and $x_{OH\text{-}COT}$ are excretion percentage of COT and OH-COT relative to NIC (%).

The average of amount of NIC consumed (m_{NIC_Mean}) was calculated by multiply each mass load estimate by the population equivalent weight (Equation 2).

$$m_{\mathrm{NIC\_Mean}} = \sum_{i=1}^{n} m_{\mathrm{NIC},i} W_{\mathrm{P},i} \quad (2)$$

where $m_{\mathrm{NIC},i}$ is mass load of WWTP of i, $W_{\mathrm{P},i}$ is the population equivalent weight for WWTP of i, n is the number of the WWTPs from 1 to 11.

The population equivalent weight for each WWTP was calculated based on the population served and total population for all 11 WWTPs (Equation 3).

$$W_{\mathrm{P},i} = \frac{P_i}{\sum_{i=1}^{n} P_i} \quad (3)$$

where P_i is the population served by the WWTP of i and $\sum_{i=1}^{n} P_i$ is the total population from 11 WWTPs.

From an epidemiological and public health perspective, a desirable goal would be to estimate the number of cigarettes consumed per smoker rather than the net mass of NIC within a community. Recognizing this, the number of cigarettes consumed per smoker (n_{C}) was calculated using Equation 4:

$$n_{\mathrm{C}} = \frac{m_{\mathrm{NIC\_Mean}}}{D \times Y \times R_{15} \times x_{\mathrm{Smoker}}} \quad (4)$$

where D is the content of NIC in an average cigarette, Y is the average yield of NIC uptake during smoking, R_{15} is the ratio of population aged 15 and older (15+) in Dalian, and x_{Smoker} is the smoker ratio (15+) in China. D was estimated at 0.90±0.15 mg of NIC per cigarette sold in Chinese market. The value of Y was determined to be 0.66±0.19 (based on the amount measured for that cigarette brand by standard Federal Trade Commission method and metabolites of NIC in urine [30]). The values of 91% and 24% for R_{15} and x_{Smoker} were used to calculate the number of cigarettes smoked per smoker.

Urinary excretion profile of NIC

Studies of NIC metabolism have been conducted over several decades. Many studies have reported excretory profiles of NIC administered through smoking, oral ingestion, and transdermal administration[31-39]. The analysis of compiled data suggests that the urinary excretion profile is quite similar when NIC is either administered through smoking or transdermal patches; however, the oral ingestion has higher extraction rates of OH-COT and COT (Table S2). The main elements of the excretory profile of NIC are the NIC metabolites COT and OH-COT (and their conjugated forms), and NIC itself [40]. The conjugated forms of these metabolites are completely transformed to the free form by *β*-glucuronidase enzymes from faecal bacteria in the raw wastewater [7]. Therefore, the total amount (percent of free and conjugated forms) of OH-COT and COT accounts for 43.4±13.8% and 32.3±8.0% of the total amount of NIC equivalents excreted when smoking is the main route of nicotine administration. These mean excretion percentages were used to calculate the mass of NIC consumed.

Uncertainty Analysis

The calculated estimates of tobacco consumption are based on the average values of the input parameters. However, their implementations are complicated since the parameters cannot be treated as fixed-point values: each parameter may take on a range of values depending on if the actual values are uncertain (e.g., concentrations of COT and OH-COT) or if they vary from person to person (e.g., excretion percentage of COT and OH-COT). The most effective quantification method for uncertainty and variability is to assign to each parameter a probability density function. The Monte Carlo approach

(Oracle Crystal Ball software, Version 7.3.1) allows the possibility of describing the uncertainty and variability associated with the input parameters and incorporating them into the estimate of tobacco consumption.

Selecting appropriate input distributions for uncertain and variable inputs is crucial to the integrity of Monte Carlo simulation results. These distributions of density functions can be taken several different forms (e.g., normal, lognormal, uniform, triangular). Normal distribution can be used to represent a quantity for which the underlying mechanism can be described by the central limit theorem, such as the resultant of a large number of additive independent errors. In this study, we assumed the flow rates, population, the content of NIC in cigarettes, and the yield of NIC uptake during smoking were normal distributions with a conservative estimate of 10% relative standard deviation. The central limit theorem can also be used as the basis for selecting a lognormal distribution to represent a quantity. The lognormal distribution has often been found to be a good representation of positive real values and results from the production of many random variation of the dilution of pollutant concentrations. As such, the concentrations and excretion percentage of COT and OH-COT were taken as lognormal distributions. Tables S3 and S4 show the description of the parameter distributions for Monte Carlo simulations.

Sensitivity analysis shows how much each input parameter contributes to the uncertainty or variability of the estimations. In turn, it adds information on which portion of the variability is from natural fluctuation versus how much is caused by lack of knowledge. The contribution of an assumption to the total variance of the forecast is

obtained as the square of the correlation coefficient normalized to the sum of the squared correlation coefficients (adjusted to 100%).

Results and discussion

Concentrations in wastewater

NIC, COT, and OH-COT were detected in all 22 samples with average concentrations of 11.5±8.73, 2.33±0.30 and 2.76±0.91 µg/L (see Table 1), respectively. As expected, OH-COT concentrations were ~1.2 times higher than those of COT, in concordance with the known metabolic fate of NIC in humans (~1.3 times). These values were highly correlated ($R^2 = 0.90$). The concentrations measured in the samples collected on 11 WWTPs were very similar and stable as expected. In contrast, parent NIC concentrations were higher than expected considering only urinary excretion, and showed variability across the different WWTPs. Cigarette butts and ash contribute to the amount of NIC in wastewater, suggesting that the human metabolites of COT and OH-COT are the best biomarkers for estimating tobacco consumption. NIC concentrations did not correlated well with either COT ($R^2 = 0.69$) or OH-COT ($R^2 = 0.72$).

These concentrations are in the same range of those measured by Lopes et al. in Lisbon, Portugal (1.1–3.5 µg/L for COT)[41] and slightly higher than those reported by Rodríguez-Álvarez et al. in Spain (2.2–5.9 µg/L, 1.0–1.6µg/L, and 1.8–2.8 µg/L for NIC, COT, and OH-COT)[6]. The concentrations can be influenced by many factors, such as tobacco prevalence and flows arriving to the WWTPs.

Table 1. The concentration of NIC, COT, and OH-COT in wastewater, flow rate, and served population, mass loads, and smoking cigarette number for 11 WWTPs in Dalian, China

WWTPs	Population ($\times 10^4$)	Flow rate ($\times 10^4$ m^3/d)	$C_{OH\text{-}COT}$ (μg/L)	C_{COT} (μg/L)	C_{NIC} (μg/L)	Mass load (mg/day/capita)	Cigarette number (cigarettes/day/smoker)
MYI	35	10	2.97±0.32	2.52±0.32	15.9±5.47	1.82	13.8
MER	25	8.15	2.99±0.33	2.55±0.32	6.81±5.84	2.10	15.9
LSH	22	5.5	2.04±0.21	1.76±0.16	4.77±1.21	1.10	8.4
FJZ	10	0.7	1.57±0.79	1.47±0.63	10.2±9.25	0.25	1.9
LHT	27	8	2.30±0.14	1.79±0.07	3.88±1.10	1.40	10.7
QSH	19.5	3.4	3.50±0.10	3.56±0.65	18.6±8.29	1.44	10.9
SEG	20	8	3.51±0.05	3.06±0.23	13.5±1.99	3.06	23.2
CYI	18.8	8	2.09±0.66	1.35±0.80	7.77±7.15	1.69	12.9
CER	28.2	12	3.22±0.30	2.97±0.06	14.7±0.62	3.07	23.3
HTX	8.1	2.5	1.77±0.28	1.03±0.22	0.67±0.74	1.00	7.6
XJH	6.5	3	4.36±0.83	3.51±0.62	29.2±3.17	4.22	32.0

Mass load of NIC consumed and uncertain analysis

The mass load of absorbed NIC during the sampling period ranges from 0.25 mg/day/capita to 4.22 mg/day/capita with an average of 1.92 mg/day/capita (Table 1). The mass load of NIC consumed reported in this work are similar than those reported in Spain (1.8 mg/day/capita) and lower than in Italy (3.42 mg/day/capita). These differences in the mass load of NIC consumed may be explained by the underlying differences in tobacco consumption.

Figure 1 shows the probability distribution density of the NIC mass load in Dalian generated by the Monte Carlo simulations. The figure shows the 10^{th} percentile, the central tendency of consumption (50^{th} percentile) and the reasonable maximum mass load (90^{th} percentile). The value of m_{NIC_Mean} was estimated to be 2.04 mg/day/capita, ranged from 1.56 mg/day/capita to 2.39 mg/day/capita in Dalian. The results reveal that the uncertainty of mass load, as defined by the ratio of the percentile 90^{th} to the percentile 10^{th}, is 1.5.

Figure 2 shows the sensitivity analysis for m_{NIC_Mean} and the eight most significant variables influencing mass load (>1%). The two most significant parameters influencing mass load were x_{COT} and $x_{OH\text{-}COT}$, which contribute 52% and 31% of total uncertainty of mass load, respectively. In contrast, population, flow rate, and concentration combined only contribute 17% of the observed variance. In this analysis, it can be seen that variability of extraction percentage are far more important than the uncertainty about population, flow rate, and concentration to overall variance in mass load. The variability of extraction percentage for smokers refers to inter-individual heterogeneity with respect

to different extraction rates of OH-COT and COT. The mechanisms underlying these differences of urinary of NIC metabolites likely involve both genetic and behavioral factors [42]. Variability is an inherent property of a system under study and therefore is irreducible. While the uncertainty of population, flow rate, and concentration due to random or systematic error in measurement can be reduced with better data and/or better models. The main uncertainty of estimation results originates from the inter-individual variability, which cannot be reduced through improving the measurement technique.

Consumption of tobacco and uncertain analysis

Given that mean value of 0.90 mg/cigarette of NIC content in Chinese market and 66% yield of NIC uptake, we estimated an average of 3.0 cigarettes/day/capita across Dalian. Assuming that 24% of population (15+) in China are active smokers, and 91% of population in Dalian are adults (15+), our procedure estimates that the average Dalian smoker consumed 14.6 CPD. This compares well to survey estimates of population smoking. The 2010 Global Adult Tobacco Survey found that adult Chinese smokers consumed, on average, 14.2 manufactured CPD [3]. Similarly, a 2009 smoking prevalence investigation among urban residents of Liaoning province—which includes Dalian—found that the smokers consumed, on average 14.2±0.3 [43]. Hence, the smoking rate obtained from sewage analysis shows good agreement with the smoking rate estimated from epidemiological survey data.

The probability distribution density for consumed number of cigarettes per smoker in Dalian is shown in Figure 3. Since this estimation is calculated as the product of several probability distributions, the mass load distribution shown in Figure 3 is approximately lognormal. The estimated cigarette consumption ranged from 10 to 27 CPD, with a

median of 16 CPD. The uncertainty and variability of number of cigarettes was estimated as 2.0. The probability distribution of daily cigarette consumption is consistent with the questionnaire review from Jiangsu Province in 2010, which showed the median number of cigarettes consumed per day was 15 and ranged from 10 to 20 for the 25th and 75th percentile [44].

The sensitivity analysis for number of cigarettes smoked shows that the four parameters of Y, D, $x_{OH\text{-}COT}$, and x_{COT} contributed to 63%, 20%, 9%, and 5% of variance, respectively (Figure 4). The contribution of the each other parameter is lower than 1%. The content of NIC in one cigarette is the most important parameters than other factors, which contribute almost half of the total uncertainty. Due to inter-individual heterogeneity of three parameters of Y, $x_{OH\text{-}COT}$, and x_{COT}, further study on improving the estimating the true values of D should be done to reduce the uncertainty of estimation for cigarette consumption.

In summary, the present study provides a systematic assessment of tobacco use in a city based on sewage epidemiology. Generally, the estimated tobacco consumption agrees quite well with those derived from a separate epidemiology survey. Monte Carlo simulation shows that the uncertainty of mass load and cigarette number are 1.5 and 2.0. The results of sensitivity analysis show that the inter-individual variability plays an important role on uncertainty of estimation results. These data are the first published estimates generated using this tool for tobacco consumption in a Chinese city. Although sewage epidemiology has some major strengths, it is not without limitations. This approach gives no information on demographic characteristics of smokers. As such, it is

unlikely that such an approach could entirely replace traditional epidemiological sampling.

Further work is required to develop new techniques to reduce the uncertainty of consumption. Sewage epidemiology is a potential approach to estimate consumption of nicotine, and could complement conventional socio-epidemiological surveys, particularly when it comes to the real-time monitoring of the impact of new tobacco control policies.

Acknowledgements

Thanks for the discussion with Dr. Lian-Zheng Yu from the Department of Noncommunicable Chronic Disease Prevention, Liaoning Center for Disease Control and Prevention.

Contributors

De-Gao Wang designed whole study and wrote the paper. Qian-Qian Dong and Juan Du pretreated the samples. Shuo Yang collected the samples. Yun-Jie Zhang designed the model. Guang-Shui Na analysed the samples. Stuart G. Ferguson provided input on nicotine metabolism and revised the paper. Zhuang Wang and Tong Zheng revised the draft paper.

Funding

This study was supported by the Fundamental Research Funds for the Central Universities, Open Fund by Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (KHK1309), A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and

Open Project of State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology (QAK201503).

Provenance and peer review

Not commissioned; externally peer reviewed.

Competing interests

None.

Data sharing statement

No additional data are available.

References

1. Li Q, Hsia J, Yang G. Prevalence of Smoking in China in 2010. New England Journal of Medicine 2011;**364**(25):2469-70.
2. Sung HY, Wang L, Jin S, et al. Economic burden of smoking in China, 2000. Tobacco Control 2006;**15**(3):i5-i11.
3. WHO. Global Adult Tobacco Survey (GATS) China 2010 Country Report.: Chinese Center for Disease Control and Prevention, 2011.
4. Wang JB, Jiang Y, Wei WQ, et al. Estimation of cancer incidence and mortality attributable to smoking in China. Cancer Causes and Control 2010;**21**(6):959-65.
5. Gartner C. Flushing out smoking: measuring population tobacco use via wastewater analysis. Tobacco Control 2015;**24**(1):1-2.
6. Rodríguez-Álvarez T, Rodil R, Rico M, et al. Assessment of local tobacco consumption by liquid chromatography–tandem mass spectrometry sewage

analysis of nicotine and its metabolites, cotinine and trans-3′-hydroxycotinine, after enzymatic deconjugation. Analytical Chemistry 2014;**86**(20):10274-81.

7. Castiglioni S, Senta I, Borsotti A, et al. A novel approach for monitoring tobacco use in local communities by wastewater analysis. Tobacco Control 2015;**24**(1):38-42.
8. Reid MJ, Langford KH, Grung M, et al. Estimation of cocaine consumption in the community: a critical comparison of the results from three complimentary techniques. Bmj Open 2012;**2**(6):1-9.
9. Li J, Hou L, Du P, et al. Estimation of amphetamine and methamphetamine uses in Beijing through sewage-based analysis. Science of the total environment 2014;**490**:724-32.
10. Du P, Li K, Li J, et al. Methamphetamine and ketamine use in major Chinese cities, a nationwide reconnaissance through sewage-based epidemiology. Water Research 2015;**84**:76-84.
11. Castiglioni S, Bijlsma L, Covaci A, et al. Evaluation of uncertainties associated with the determination of community drug use through the measurement of sewage drug biomarkers. Environmental Science and Technology 2012;**47**(3):1452-60.
12. Reid MJ, Langford KH, Mørland J, et al. Quantitative assessment of time dependent drug-use trends by the analysis of drugs and related metabolites in raw sewage. Drug and Alcohol Dependence 2011;**119**(3):179-86.
13. van Nuijs ALN, Castiglioni S, Tarcomnicu I, et al. Illicit drug consumption estimations derived from wastewater analysis: A critical review. Science of The Total Environment 2011;**409**:3564-77.

14. van Nuijs ALN, Mougel J-F, Tarcomnicu I, et al. Sewage epidemiology - A real-time approach to estimate the consumption of illicit drugs in Brussels, Belgium. Environment International 2011;**37**(3):612-21.

15. Postigo C, Lopez de Alda MJ, Barceló D. Analysis of drugs of abuse and their human metabolites in water by LC-MS^2: A non-intrusive tool for drug abuse estimation at the community level. TrAC Trends in Analytical Chemistry 2008;**27**(11):1053-69.

16. Ort C, van Nuijs AL, Berset JD, et al. Spatial differences and temporal changes in illicit drug use in Europe quantified by wastewater analysis. Addiction 2014;**109**(8):1338-52.

17. Thomas KV, Bijlsma L, Castiglioni S, et al. Comparing illicit drug use in 19 European cities through sewage analysis. Science of The Total Environment 2012;**432**:432-9.

18. Lai FY, Ort C, Gartner C, et al. Refining the estimation of illicit drug consumptions from wastewater analysis: Co-analysis of prescription pharmaceuticals and uncertainty assessment. Water Research 2011;**45**(15):4437-48.

19. Mathieu C, Rieckermann J, Berset J-D, et al. Assessment of total uncertainty in cocaine and benzoylecgonine wastewater load measurements. Water Research 2011;**45**(20):6650-60.

20. Castiglioni S, Bijlsma L, Covaci A, et al. Evaluation of uncertainties associated with the determination of community drug use through the measurement of sewage drug biomarkers. Environmental Science and Technology 2013;**47**(3):1452-60.

21. Ort C, Eppler JM, Scheidegger A, et al. Challenges of surveying wastewater drug loads of small populations and generalizable aspects on optimizing monitoring design. Addiction 2014;**109**(3):472-81.

22. Ort C, Lawrence MG, Rieckermann J, et al. Sampling for pharmaceuticals and personal care products (PPCPs) and illicit drugs in wastewater systems: are your conclusions valid? A critical review. Environmental Science and Technology 2010;**44**(16):6024-35.

23. Thai PK, Jiang G, Gernjak W, et al. Effects of sewer conditions on the degradation of selected illicit drug residues in wastewater. Water Research 2014;**48**:538-47.

24. Chen C, Kostakis C, Gerber JP, et al. Towards finding a population biomarker for wastewater epidemiology studies. Science of the Total Environment 2014;**487**:621-28.

25. Daughton CG. Real-time estimation of small-area populations with human biomarkers in sewage. Science Of the Total Environment 2012;**414**:6-21.

26. Lai FY, Anuj S, Bruno R, et al. Systematic and day-to-day effects of chemical-derived population estimates on wastewater-based drug epidemiology. Environmental Science and Technology 2014;**49**(2):999-1008.

27. Khan U, Nicell JA. Refined sewer epidemiology mass balances and their application to heroin, cocaine and ecstasy. Environ Int 2011;**37**(7):1236-52.

28. Khan U, Nicell JA. Sewer epidemiology mass balances for assessing the illicit use of methamphetamine, amphetamine and tetrahydrocannabinol. Science of the Total Environment 2012;**421**:144-62.

29. Brewer AJ, Ort C, Banta-Green CJ, et al. Normalized diurnal and between-day trends in illicit and legal drug loads that account for changes in population. Environmental Science and Technology 2012;**46**(15):8305-14.

30. St.Charles FK, Kabbani AA, Borgerding MF. Estimating tar and nicotine exposure: Human smoking versus machine generated smoke yields. Regulatory Toxicology and Pharmacology 2010;**56**(1):100-10.

31. Byrd GD, Chang KM, Greene JM, et al. Evidence for urinary excretion of glucuronide conjugates of nicotine, cotinine, and trans-3'-hydroxycotinine in smokers. Drug Metabolism and Disposition 1992;**20**(2):192-7.

32. Gunilla A, Gunilla B, Margareta C. Oral mucosal changes and nicotine disposition in users of Swedish smokeless tobacco products: a comparative study. Journal of Oral Pathology and Medicine 1994;**23**(4):161-67.

33. Byrd GD, Robinson JH, Caldwell WS, et al. Comparison of measured and FTC-predicted nicotine uptake in smokers. Psychopharmacology 1995;**122**(2):95-103.

34. Andersson G, Vala EK, Curvall M. The influence of cigarette consumption and smoking machine yields of tar and nicotine on the nicotine uptake and oral mucosal lesions in smokers. Journal of Oral Pathology and Medicine 1997;**26**(3):117-23.

35. Byrd GD, Davis RA, Caldwell WS, et al. A further study of FTC yield and nicotine absorption in smokers. Psychopharmacology 1998;**139**(4):291-99.

36. Meger M, Meger-Kossien I, Schuler-Metz A, et al. Simultaneous determination of nicotine and eight nicotine metabolites in urine of smokers using liquid

chromatography–tandem mass spectrometry. Journal of Chromatography B 2002;**778**(1-2):251–61.

37. Benowitz NL, Pomerleau OF, Pomerleau CS, et al. Nicotine metabolite ratio as a predictor of cigarette consumption. Nicotine and Tobacco Research 2003;**5**(5):621-24.

38. Mcguffey JE, Wei B, Bernert JT, et al. Validation of a LC-MS/MS method for quantifying urinary nicotine, six nicotine metabolites and the minor tobacco alkaloids--anatabine and anabasine--in smokers' urine. Plos One 2014;**9**(9).

39. Benowitz NL, Fong I, Gupta S. Nicotine metabolic profile in man: comparison of cigarette smoking and transdermal nicotine. Journal of Pharmacology and Experimental Therapeutics 1994;**268**(1):296-303.

40. Jacobson GA, Ferguson SG. Relationship between cotinine and trans-3′-hydroxycotinine glucuronidation and the nicotine metabolite ratio in Caucasian smokers. Biomarkers 2014;**19**(8):679-83.

41. Lopes A, Silva N, Bronze MR, et al. Analysis of cocaine and nicotine metabolites in wastewater by liquid chromatography-tandem mass spectrometry. Cross abuse index patterns on a major community. Science of the Total Environment 2014;**487**:673-80.

42. Derby K, Cuthrell K, C, Carmella S, et al. Nicotine metabolism in three ethnic/racial groups with different risks of lung cancer. Cancer Epidemiology, Biomarkers and Prevention 2008;**17**(12):3526-35.

43. Yu L, Mei D, An X, et al. Analysis on diffierence in smoking prevalence among urban residents in Liaoning Province, 1999-2009. Chinese Journal of Disease Control and Prevention 2012;**16**(10):842-45.

44. Su J, QIn Y, Xiang Q, et al. Smoking behaviors and tobacco consumption of smokers among residents in Jiansu Province. Chinese Journal of Public Health 2014;**30**(11):1368-72.

Figure captions

Figure 1. Distribution of estimated mass load of NIC (mg/day/capita) in Dalian: mean, median, 10th, and 90th percentiles.

Figure 2. The results of sensitivity analysis for NIC mass load. x_{COT} and $x_{OH\text{-}COT}$ are excretion percentage of COT and OH-COT relative to NIC (%); P_{cer} and F_{cer} are the population and flow rate in WWTP of CER; C_{COT_cyi} is the concentration of COT in WWTP of CYI; P_{seg} and F_{seg} are the population and flow rate in WWTP of SEG; P_{myi} is the population in WWTP of MYI.

Figure 3. Distribution of estimated number of cigarette consumed (cigarettes/day/smoker) in Dalian: mean, median, 10th and 90th percentiles.

Figure 4. The results of sensitivity analysis for number of cigarettes consumed. x_{COT} and $x_{OH\text{-}COT}$ are excretion percentage of COT and OH-COT relative to NIC (%); Y is the yield of NIC during smoking, and D is the content of NIC in one cigarette.

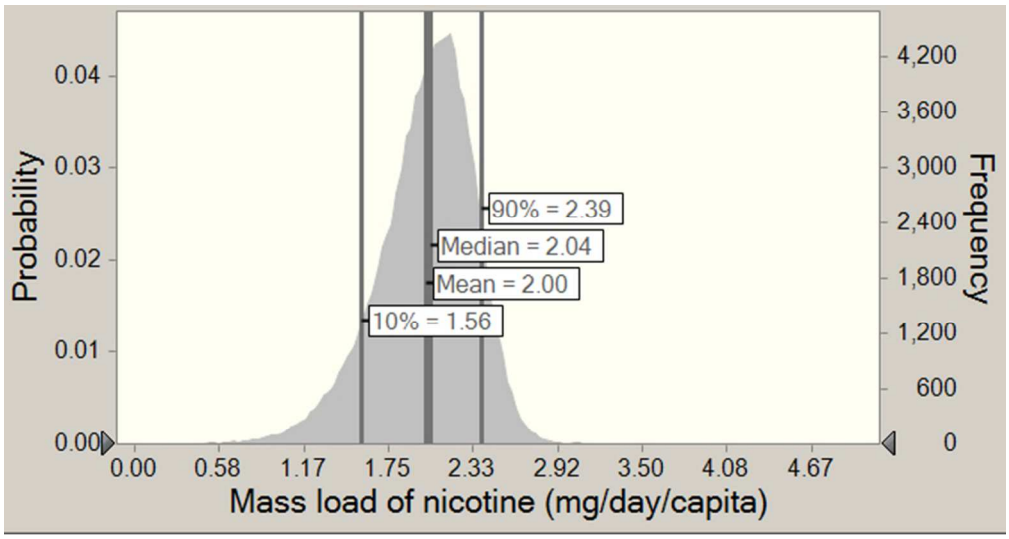

Figure 1. Distribution of estimated mass load of NIC (mg/day/capita) in Dalian: mean, median, 10th, and 90th percentiles.
56x29mm (300 x 300 DPI)

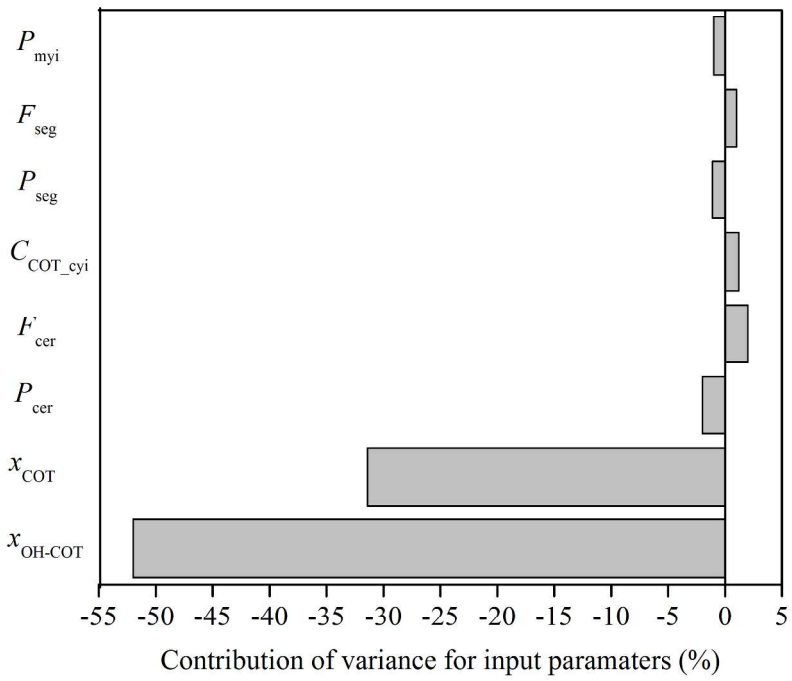

Figure 2. The results of sensitivity analysis for NIC mass load. xCOT and xOH-COT are excretion percentage of COT and OH-COT relative to NIC (%); Pcer and Fcer are the population and flow rate in WWTP of CER; CCOT_cyi is the concentration of COT in WWTP of CYI; Pseg and Fseg are the population and flow rate in WWTP of SEG; Pmyi is the population in WWTP of MYI.
254x203mm (300 x 300 DPI)

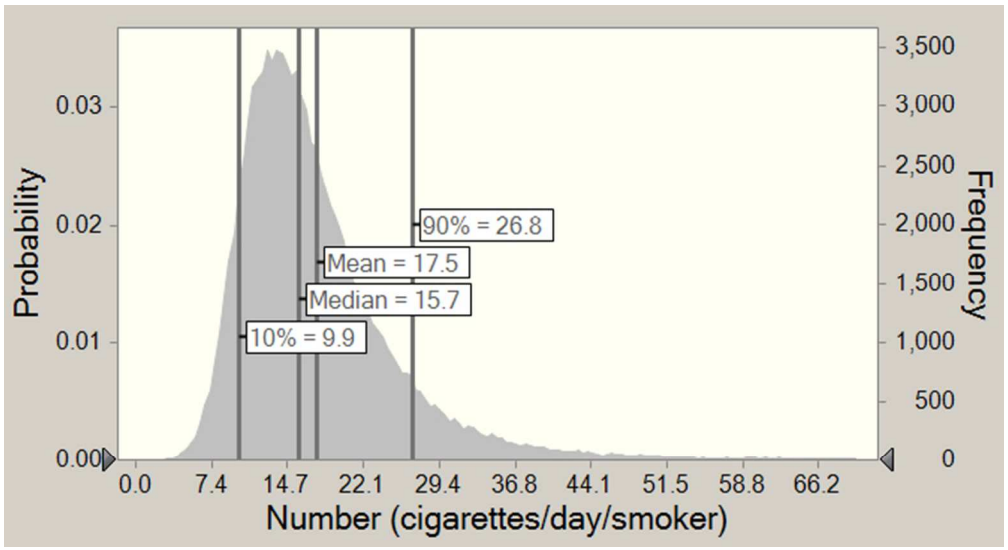

Figure 3. Distribution of estimated number of cigarette consumed (cigarettes/day/smoker) in Dalian: mean, median, 10th and 90th percentiles.
56x30mm (300 x 300 DPI)

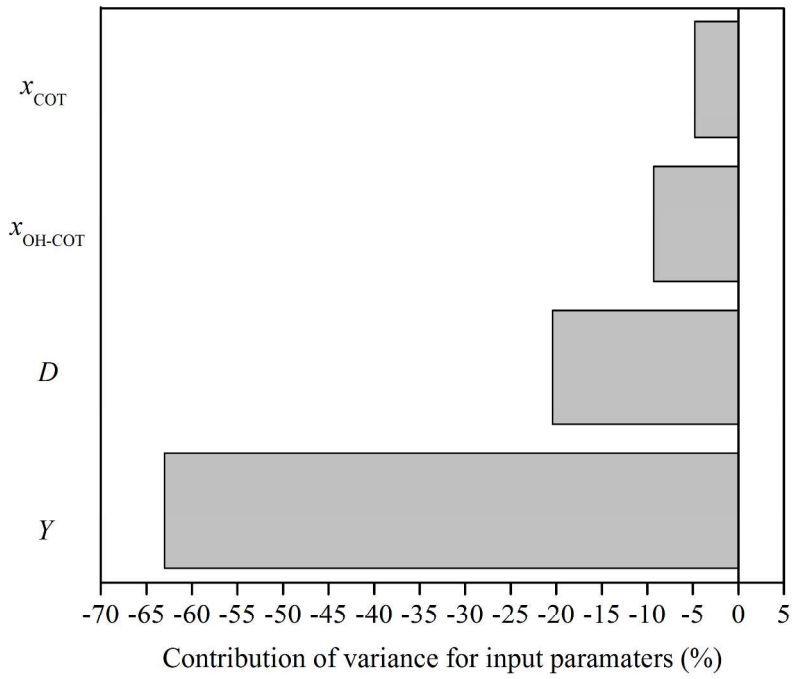

Figure 4. The results of sensitivity analysis for number of cigarettes consumed. xCOT and xOH-COT are excretion percentage of COT and OH-COT relative to NIC (%);Y is the yield of NIC during smoking, and D is the content of NIC in one cigarette.
254x203mm (300 x 300 DPI)

Supporting Information

Monte Carlo simulation assess variability and uncertainty of tobacco consumption in a city by sewage epidemiology

De-Gao Wang[1†*], Qian-Qian Dong[1], Juan Du[1], Shuo Yang[1], Yun-Jie Zhang[2], Guang-Shui Na[3], Stuart G. Ferguson[4], Zhuang Wang[5], Tong Zheng[6]

[1]School of Environmental Science and Technology, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[2]Department of Mathematics, Dalian Maritime University, 1 Linghai Road, Dalian, Liaoning, China, 116023

[3]National Marine Environmental Monitoring Center, Dalian, Liaoning, China, 116023

[4]Faculty of Health, School of Medicine, University of Tasmania Private Bag 34, Hobart TAS 7000

[5]Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China

[6]State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology, Harbin 150090

*Corresponding author phone: 86-411 8472-3185; fax: 86-411 8472-3185

E-mail address: degaowang@dlmu.edu.cn (De-Gao Wang)

Chemicals

NIC and cotinine-d3 (COT-d3) were purchased from Sigma−Aldrich Inc. (Saint Louis, MO, U.S.A). Nicotine-d3 (NIC-d3) was purchased from Toronto Research Chemicals Inc. (TRC, ON, Canada). COT was purchased from ChromaDex, Inc. (ChromaDex, CA, USA). OH-COT and trans-3′-hydroxycotinine-d3 (OH-COT-d3) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Individual stock solutions were prepared in acetonitrile and diluted as necessary. HPLC-grade methanol and acetonitrile (ACN) were purchased from J&K Chemical (J&K, China). Ultrapure water was obtained from a Milli-Q water generator (Millipore, USA).

Analysis

Raw wastewater (20 mL) was passed through a glass filter (0.45 μm) to removal particles. Then, deuterated labeled internal standards (20 ng) for quantification were added and a solid phase extraction (SPE) clean-up set with reverse-phase cartridges (Cleanert PEP-2, 1 mL, 60 mg, China) was applied. The cartridges were conditioned with consecutively methanol (2 mL) and Milli-Q water (2 mL). The extracts were allowed to enter the preconditioned SPE cartridge connected to a vacuum pump. The cartridge was washed with 1 mL of Milli-Q water. After drying of the cartridges, the elution was performed with 1 mL of methanol. The eluate was evaporated and redissolved in 1 mL of ACN for final analysis.

The liquid chromatography coupled with tandem mass-spectrometry (DSQ600, Canada) (LC–MS/MS) was used to analyze the samples. The injection volume was 10 μL. Targeted analytes were chromatographically separated using a Waters Symmetry Shield

C18 column (3.5 μm, 150 mm×2.1 mm, Waters, USA) at a flow rate of 0.4 mL/min. Eluents used for the separation included 5 mmol/L NH_4AC in H_2O (eluent A) and 5 mmol/L NH_4AC in methanol (eluent B). The gradient used as follows: Start as 0 min with 20% B and 80% A, held for 2 min; 2-6 min linear rate to 100% B, held for 2 min; 8-10 min linear rate to 20% B, 10-11 min return to initial conditions; 20% B.

Scheduled multiple reaction monitoring with positive electrospray ionisation were operated to identify and quantify the masses of analytes. The collision energy and tube lens offset were optimised for each analyte and surrogate separately. Positive ionization electrospray was used with nitrogen as the cone and desolvation gas. Identification and quantification was performed using two characteristic transitions for the analysed compounds (Table S1).

Quantification was performed using a seven-point calibration curve spanning the range of anticipated analyte concentrations in the samples. The method precision and recovery were tested for both matrices, by spiking three parallel samples of Milli-Q water and wastewater with a mixture of all analytes at environmentally relevant concentrations. The recoveries of analytes varied between 64% and 93%. All concentrations of analytes were corrected by the recoveries. The limits of detection were estimated to be 0.020 μg/L, 0.012 μg/L, and 0.014 μg/L for NIC, COT, and OH-COT, respectively.

Table S1 Target compounds and stable isotope standards with the retention times, the mass spectrometric SRM ion transitions, de-clustering potential (DP) and collision energies (CE) for analysis, and the limits of detection (LOD) and recoveries for target analytes

Compound	Retention Time	SRM Transition	DP(V)	CE(V)	LOD (μg/L)	Recovery (%)
NIC	2.96	163→117	84	35	0.020	64
NIC-d3	2.96	166→106	48	22	-	-
COT	1.67	177→80	60	27	0.012	72
COT-d3	1.67	180→80	60	28	-	-
OH−COT	1.60	193→80	56	29	0.014	93
OH−COT-d3	1.60	196→80	56	48	-	-

Table S2 Distribution of urinary nicotine metabolites as percentage of the total amount of metabolites excreted (mean and standard deviation)

Reference	Route	Dose (mg)	n	Age (yr)	Duration (h)	OH-COT	OH-COT-G	COT	COT-G	NIC	NIC-G	OH-COT-T	COT-T	NIC-T
Byrd et al.(1992)	Smoked	7.2-48	11	29-47	24	35±8	9±4	13±4	17±7	10±4	3±2	45.3±10.4	29.8±7.8	12.7±3.6
Benowitz et al. (1994)	Smoked	34.2±21.4	12		24	39.1±12.5	7.8±5.9	13.3±3.1	15.8±7.8	10.4±4.4	4.6±2.4	46.9±18.4	29.1±10.9	15.0±7.3
Byrd et al.(1995)	Smoked	4.8-38.1	33	26-50	24	-	-	-	-	-	-	53.3±15.6	26.6±10.8	10.4±3.9
Andersson et al. (1997)	Smoked	25.0±10.7	91	38.0±12.1	24	36.1±10.6	22.8±10.0	9.2±2.6	14.0±5.4	9.4±5.7	4.5±2.5	58.9±20.6	23.2±8.0	13.9±8.2
Byrd et al.(1998)	Smoked	34±8	72	37±9	24	-	-	-	-	-	-	43.1±12.2	25.4±7.1	14.4±6.8
Meger et al.(2002)	Smoked	-	5	-	24	32.2±14	6.8±3.2	16.1±2.1	19.6±6.9	10.4±4.1	4.5±4	39.0±17.2	35.7±9.0	14.9±8.1
Benowitz et al. (2002)	Smoked	12.3±10.4	40	7.7±6.2	24	23.5±16.3	8.9±8.6	26.0±10.3	12.1±9.1	19.2±16.3	10.3±7.3	32.4±24.9	38.1±19.4	29.5±23.6
McGuffey et al. (2014)	Smoked	-	94	-	24	-	-	-	-	-	-	38.8±20.7	26.4±9.6	20.6±14.2
Average	Smoked	-	154	-	24	-	-	-	-	-	-	**43.4±13.8**	**32.3±8.0**	**19.4±8.7**
Benowitz et al. (1994)	Transdermal	21.5±4.5	12	-	24	37.0±10.8	7.9±4.7	14.9±4.6	15.4±7.9	11.1±4.3	5.3±3.3	44.9±15.5	30.3±12.5	16.4±7.6
Andersson et al. (1994)	Oral	9.3-124	54	43.3±10.7	24	41.6±10.6	19.4±11.0	7.9±2.2	8.9±4.6	8.3±5.7	3.0±1.8	61.0±21.6	16.8±6.8	11.3±7.5

Table S3. Monte Carlo parameter distribution for population, flow rate, and concentrations in 11 WWTP for estimation of tobacco use

WWTP	Symbol	Concentrations (μg/L)					Flow rate ($\times 10^4$ m^3/d)			Served population of WWTP ($\times 10^4$)			Weight		
		Type	Distribution	Symbol	Type	Distribution	Symbol	Type	Distribution	Symbol	Type	Distribution	Symbol	Type	Distribution
MYI	$C_{OH\text{-}COT\_myi}$	lognormal	2.97±0.32	C_{COT_myi}	lognormal	2.52±0.32	F_{myi}	normal	10±1.0	P_{myi}	normal	35±3.5	W_{myi}	point	0.22
MER	$C_{OH\text{-}COT\_mer}$	lognormal	2.99±0.33	C_{COT_mer}	lognormal	2.55±0.32	F_{mer}	normal	8.15±0.82	P_{mer}	normal	25±2.5	W_{mer}	point	0.18
LSH	$C_{OH\text{-}COT\_lsh}$	lognormal	2.04±0.21	C_{COT_lsh}	lognormal	1.76±0.16	F_{lsh}	normal	5.5±0.55	P_{lsh}	normal	22±2.2	W_{lsh}	point	0.08
FJZ	$C_{OH\text{-}COT\_fjz}$	lognormal	1.57±0.79	C_{COT_fjz}	lognormal	1.47±0.63	F_{fjz}	normal	0.7±0.07	P_{fjz}	normal	10±1.0	W_{fjz}	point	0.01
LHT	$C_{OH\text{-}COT\_lht}$	lognormal	2.30±0.14	C_{COT_lht}	lognormal	1.79±0.07	F_{lht}	normal	8±0.8	P_{lht}	normal	27±2.7	W_{lht}	point	0.13
QSH	$C_{OH\text{-}COT\_qsh}$	lognormal	3.50±0.10	C_{COT_qsh}	lognormal	3.56±0.65	F_{qsh}	normal	3.4±0.34	P_{qsh}	normal	19.5±1.95	W_{qsh}	point	0.10
SEG	$C_{OH\text{-}COT\_seg}$	lognormal	3.51±0.05	C_{COT_seg}	lognormal	3.06±0.23	F_{seg}	normal	8±0.8	P_{seg}	normal	20±2.0	W_{seg}	point	0.21
CYI	$C_{OH\text{-}COT\_cyi}$	lognormal	2.09±0.66	C_{COT_cyi}	lognormal	1.35±0.80	F_{cyi}	normal	8±0.8	P_{cyi}	normal	18.8±1.88	W_{cyi}	point	0.11
CER	$C_{OH\text{-}COT\_cer}$	lognormal	3.22±0.30	C_{COT_cer}	lognormal	2.97±0.06	F_{cer}	normal	12±1.2	P_{cer}	normal	28.2±2.82	W_{cer}	point	0.30
HTX	$C_{OH\text{-}COT\_htx}$	lognormal	1.77±0.28	C_{COT_htx}	lognormal	1.03±0.22	F_{htx}	normal	2.5±0.25	P_{htx}	normal	8.1±0.81	W_{htx}	point	0.03
XJH	$C_{OH\text{-}COT\_xjh}$	lognormal	4.36±0.83	C_{COT_xjh}	lognormal	3.51±0.62	F_{xjh}	normal	3±0.3	P_{xjh}	normal	6.5±0.65	W_{xjh}	point	0.09

Table S4. Monte Carlo parameter distribution of other input parameters for estimation of tobacco use

Parameter	Symbol	Units	Type	Distribution
Molecular weight of NIC	M_{NIC}	unitless	point	162
Molecular weight of COT	M_{COT}	unitless	point	176
Molecular weight of OH-COT	$M_{OH\text{-}COT}$	unitless	point	192
Excretion percentage of COT	x_{COT}	unitless	lognormal	32.3%±8.0%
Excretion percentage of OH-COT	$x_{OH\text{-}COT}$	unitless	lognormal	43.4%±13.8%
Population ratio in Dalian (>15)	R_{15}	unitless	point	0.91
Smoker ratio in China (>15)	X_{Smoker}	unitless	point	0.24
Content of NIC in one cigarette	D	mg	normal	0.90±0.15
Yield of NIC uptake	Y	unitless	normal	0.66±0.19

Research checklist

NO.