Environmental Research 166 (2018) 344–347

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Short communication

“How do you know those particles are from cigarettes?”: An algorithm to help differentiate second-hand tobacco smoke from background sources of household fine particulate matter

Ruaraidh Dobson*, Sean Semple

Institute for Social Marketing, University of Stirling, United Kingdom

ARTICLE INFO

Keywords:
Second-hand smoke
Tobacco smoke exposure
Air quality monitoring
Particulate matter

ABSTRACT

Background: Second-hand smoke (SHS) at home is a target for public health interventions, such as air quality feedback interventions using low-cost particle monitors. However, these monitors also detect fine particles generated from non-SHS sources.

The Dylos DC1700 reports particle counts in the coarse and fine size ranges. As tobacco smoke produces far more fine particles than coarse ones, and tobacco is generally the greatest source of particulate pollution in a smoking home, the ratio of coarse to fine particles may provide a useful method to identify the presence of SHS in homes.

Methods: An algorithm was developed to differentiate smoking from smoke-free homes. Particle concentration data from 116 smoking homes and 25 non-smoking homes were used to test this algorithm.

Results: The algorithm correctly classified the smoking status of 135 of the 141 homes (96%), comparing favourably with a test of mean mass concentration.

Conclusions: Applying this algorithm to Dylos particle count measurements may help identify the presence of SHS in homes or other indoor environments. Future research should adapt it to detect individual smoking periods within a 24 h or longer measurement period.

1. Introduction

Second-hand smoke (SHS) is a serious cause of poor indoor air quality in homes. Around 40% of children are regularly exposed worldwide,(GTSS Collaborative Group, 2006) putting them at risk of serious illness and impaired lung development (US Surgeon General, 2006).

For that reason interventions to promote smoke-free homes are of significant public health interest. Several interventions have been developed using air quality monitoring to inform parents of the impact of smoking on their indoor air quality, and the consequent effects on their children.(Dobson et al., 2017; Klepeis et al., 2013; Rosen et al., 2015; Wilson et al., 2013) A low-cost air quality monitor, the Dylos DC1700, has proved useful for monitoring $PM_{2.5}$ as a proxy for SHS in smokers' homes in these kinds of interventions.(Semple et al., 2015, 2013) The Dylos is a small, portable monitor which provides comparable accuracy at a considerably lower price than other widely used optical particle counters, such as the TSI Sidepak. In addition to being approximately one-tenth of the cost of the Sidepak instrument, the Dylos has several specific advantages in terms of low noise, simplicity of use and the ability to determine particle size distribution in terms of fine and coarse particulate (Semple et al., 2013)

$PM_{2.5}$ has been widely used as a proxy to quantify indoor concentrations of SHS in many settings including bars, homes and vehicles (Apelberg et al., 2013; Gorini et al., 2005) as reliable measurements can be taken easily and affordably over time using optical particle counters, in contrast to the high cost and complexity of more specific methods such as air nicotine measurement. Other activities in these settings can generate $PM_{2.5}$. These can include cooking emissions, combustion such as candle burning or the use of solid fuels for heating, and aerosols such as deodorants and hair sprays.(He et al., 2004) These sources can produce high concentrations of PM within a home which could be confused for SHS in interpretation.

Parents in previous intervention trials have been observed to deny and challenge messages about the risk of SHS, (Passey et al., 2016) and if feedback wrongly identifies non-SHS sources as being smoking activity this is likely to weaken the effectiveness of such approaches and make the participant question the validity of the measurement method. Developing reliable and accurate information on PM concentrations that are specifically linked to SHS is therefore important in the

* Corresponding author.
E-mail address: r.p.dobson@stir.ac.uk (R. Dobson).

https://doi.org/10.1016/j.envres.2018.06.019
Received 19 January 2018; Received in revised form 16 May 2018; Accepted 11 June 2018
0013-9351/

development of effective interventions.

The particle size distribution of tobacco smoke is known to skew towards fine and ultrafine particles. (Klepeis et al., 2003) The mean diameter of particles in tobacco smoke has been measured as 0.27 µm (in the case of mainstream smoke) and 0.09 µm (for sidestream smoke); smaller mean diameters than those associated with common household activities like frying, cleaning and the movement of people, and other sources (Abt et al., 2000) while still producing a sustained increase in particle mass concentration over time.(Semple and Latif, 2014)

The Dylos DC1700 provides data on both the fine and coarse fractions of particulate matter in the form of particle counts for particles larger than 0.5 µm and particles larger than 2.5 µm. It may therefore be possible to use this particle size information to distinguish between different sources of PM in a home, and potentially to classify homes as smoking or non-smoking.

This research uses particle concentration data measured in homes to develop and test a rule-based approach to determine whether tobacco was smoked in the home during the monitoring period. This information could be useful in providing air quality data to support behavioural interventions designed to encourage smokers to keep their homes smoke-free.

2. Materials and methods

2.1. Measuring mass concentrations and particle counts in homes

Previously reported methods (Semple et al., 2013) were used to assess $PM_{2.5}$ concentrations in homes. From previous work by our group (Semple, 2016), time resolved $PM_{2.5}$ data were already available from 116 smoking homes. Data from non-smoking homes were collected in the course of this research. Minute-by-minute particle counts reported by the Dylos DC1700 monitor were converted to estimated $PM_{2.5}$ concentrations using a previously developed equation (Equation 1).

$$PM_{2.5} = 0.65 + 4.16 \times 10^{-5}(\textit{Dylos total particle count} - \textit{large particle count}) + 1.57 \times 10^{-11}(\textit{Dylos total particle count} - \textit{large particle count})^2$$

Equation 1 - Conversion of Dylos particle counts to approximate mass concentration (Semple et al., 2013)

Also, the large particle percentage, consisting of the particles larger than 2.5 µm as a percentage of the total particles detected, was calculated for each minute for use in the algorithm.

2.2. Algorithm development

A four-step algorithm was developed to classify homes as smoking or non-smoking based on one day or more of Dylos-recorded data by excluding data points which were unlikely to be related to smoking. This algorithm was designed to use the ratio of large to small particles detected by the Dylos as a “signature” for the presence of SHS. Additional steps were intended to reduce noise in the data caused by brief fluctuations in levels of PM.

For each home:

1. Remove data where $PM_{2.5}$ concentration is below 5 µg/m^3. This step is intended to account for low ambient concentrations of $PM_{2.5}$ which are not related to SHS. 5 µg/m^3 was chosen as indoor $PM_{2.5}$ has been shown to correlate to 79% of ambient $PM_{2.5}$ in similar conditions (Cyrys et al., 2004), while the average ambient $PM_{2.5}$ concentration in Scotland has been modelled at 6.6 µg/m^3.(Sykes, 2016) Previous research on smoke-free homes has shown
2. For each minute of data, calculate the percentage of the total detected particles which are larger than 2.5 µm in diameter. Remove data where the percentage of large particles is greater than a threshold (described throughout as the ‘Large Particle Threshold’ or LPT).
3. Remove data where a peak lasts for fewer than three minutes, to account for random fluctuations compared to the sustained impact of SHS on indoor air quality.(Semple and Latif, 2014)
4. Take the percentage of minutes in the log where data has not been removed in one of the steps above. This can be used as an “SHS score” to classify the home as smoking or non-smoking if the score is above a cut-off (determined experimentally).

2.3. Statistical analysis

Use of the algorithm relies on two factors: the LPT which best indicates smoking, and the best-performing cut-off value for the SHS score, over which a log can be classified as smoking. Receiver operating characteristic curves were used to determine these factors.

ROC curves are a common method for determining the efficacy of a diagnostic test. (Bewick et al., 2004) In an ROC curve, a test is carried out on a set of records, and its specificity and selectivity are plotted. This allows comparison between different tests using the area under the curve (AUC) of this plot – a mathematical representation of the overall effectiveness of the test. Tests which classify records more successfully than random have AUC values greater than 0.5, while a hypothetical perfect test would have a value of 1.0.

Variants of the algorithm using LPTs between 0.1% and 4.0% (stepped up in 0.1% increments) were applied to the full dataset of logs and the categorisation results plotted on an ROC curve using IBM SPSS v24.(IBM Corp, 2016) The LPT which resulted in the highest AUC was selected, and the curve analysed to find the SHS score cut-off which maximised selectivity and specificity. An ROC curve was also generated using the mean $PM_{2.5}$measured in each household as a predictor of smoking status. Custom Python 2.7 scripts were developed to apply the algorithm to Dylos data logs.

2.4. Smoke-free homes data collection

Participants working at three health charities in Scotland were recruited. Only people living in homes where smoking or e-cigarette use was not permitted were eligible to participate in the study. A target of 30 people was set as achievable with the time and resources available.

Participants were given a Dylos DC1700 monitor and an instruction sheet asking them to install and run the monitor for 48 h in their main living space, elevated above floor level and away from doors and windows. This mirrored instructions given during previous studies of personal exposure to SHS.(Semple et al., 2012) Participants were also asked to keep a diary of events which could cause elevated PM in the home, including cooking and heating use.

After the monitoring period, the Dylos was returned to the research team and data was downloaded from it. A short report on air quality in the home was prepared for the participant and emailed to them, along with any relevant information on reducing air pollution in their home. The monitor's memory was then cleared prior to use with the next participant.

2.5. Smoking homes data

The pre-existing smoking homes dataset comprised minute-by-minute measurements from 116 homes, each spanning approximately 5 days, taken from the First Steps 2 Smoke-free (FS2SF) study (Semple, 2016). Participants in that study self-reported that smoking took place regularly within the home. No data on other events which could affect air quality was available from these homes.

2.6. Ethics

Ethical approval for this study was given by the College Ethics Review Board of the College of Life Sciences and Medicine at the University of Aberdeen.