

Anthropogenic and geogenic impacts on arsenic bioaccessibility in UK topsoils.

J D Appleton, M R Cave and J Wragg

British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

* Corresponding author: J. D. Appleton, British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

Tel: +44 (0) 115 936 3207

Fax: +44 (0) 115 936 3200

E-mail address: jda@bgs.ac.uk (J. D. Appleton)

Abstract

Predictive linear regression (LR) modelling between bioaccessible arsenic (B-As) and a range of total elemental compositions and soil properties was executed in order to assess the potential for developing a national B-As dataset for the UK. LR indicates that total arsenic (As) is the only highly significant independent variable for estimating B-As in urban areas where it explains 75-92% of the variance. The broad compatibility of the London, Glasgow and Swansea regression models suggests that application of these models to estimate bioaccessible As in UK soils impacted by diffuse anthropogenic urban contamination and non-ferrous metal processing should be relatively accurate. In areas dominated by Jurassic ironstones and associated clays and limestones, total As, P and pH are significant, accounting for 53, 14 and 5%, respectively, of the B-As variance. Models based on total As as the sole predictor in the combined Jurassic and Cretaceous sedimentary ironstones datasets explain about 40% of the B-As variance. The median As bioaccessible fraction (%As-BAF) is 19 to 28% in the anthropogenic contamination impacted urban domains, but much lower (5-9%) in geogenic terrains dominated by ironstones. Results of this study can be used as part of a lines of evidence approach to localised risk assessment but should not be used to replace bioaccessibility testing at individual sites where local conditions may vary considerably from the broad overview presented in this study.

Keywords: bioaccessibility, arsenic, geogenic, anthropogenic, topsoil, UK

1. Introduction

A substantial proportion of the landscape of the UK has naturally elevated topsoil concentrations of total arsenic (As) which exceed the Soil Guideline Value (SGV), a threshold used in the preliminary assessments for land contamination (DEFRA-EA, 2002). Inorganic As SGVs for residential, allotment and commercial land uses are 32, 43 and 640 mg kg⁻¹, respectively (Martin et al., 2009). Approximately 1300 km² of the land surface of England underlain by ironstone and about 2200 km² of land impacted by mineralisation have topsoil As >32 mg kg⁻¹ (Ander et al., 2012).

The main exposure pathway for As in soil is via soil ingestion (Paustenbach, 2000; DEFRA 2002; Bacigalupo and Hale, 2012), therefore, from a human health perspective, it is not the total amount of As in the soil but the fraction that is absorbed into the body during soil ingestion *i.e.* the bioavailable fraction, that is important for assessing human health risk. The measurement of the bioavailability of As in soil requires *in vivo* testing using humans or animal surrogates, a time consuming, costly and ethically challenging process. However, *in vitro* bioaccessibility testing has been developed and validated specifically to provide a conservative estimate of bioavailability (Basta et al., 2007; Caboche, 2009; Denys et al., 2012; Juhasz et al., 2007a). The *in vitro* bioaccessibility assay estimates the fraction of As released from the soil into solution in the gastro-intestinal (GI) tract in a form that can potentially be absorbed into the blood stream (Paustenbach, 2000; Wragg and Cave, 2003; Intawongse and Dean, 2006). Guidelines for the use of data produced by *in vitro* bioaccessibility testing methods in human health risk assessment have recently been produced in order to assist the risk assessment and regulatory communities (Nathanail, 2009).

The bioaccessibility and hence bioavailability of any contaminant bound to the soil depends upon the soil type, properties of the soil, particle size, the contaminant and the manner by which the contaminant has entered the soil (Palumbo-Roe et al., 2005; Selinus, 2005; Juhasz et al., 2007a,b; Wragg et al., 2007; Meunier et al., 2011a,b; Mingot et al., 2011). Cave et al. (2011) have described how specific properties of soil, such as pH, organic matter content, mineral constituents, solid phase partitioning of As and soil ageing may influence bioaccessibility of As and other contaminants. A number of workers have successfully carried out multiple linear regression (MLR) modelling of the bioaccessible As content of soils using their physico-chemical properties, such as the elemental composition of the soil and soil pH as the predictor variables (Yang et al., 2002; Klinck et al., 2005; Tang et al., 2007; Juhasz et al., 2007b, 2008). Sarkar et al. (2007) used the element composition, pH, clay content and cation exchange capacity and Cave et al. (2007) used spectral properties derived from near infra-red spectra of the soils. If a model is robust it can be used to predict bioaccessibility from soil properties so that the *in vitro* bioaccessibility test does not have to be carried out on every soil from a given soil region. However, site specific investigations would have to be carried out. In addition, the models predictor variables and the relative size of their coefficients can provide an insight into the processes governing the bioaccessibility of As in the soils. However, the model is useful only if the soil properties used for prediction of bioaccessibility are already known or are more easily and cheaply measured than the *in vitro* bioaccessibility test. Furthermore, the MLR models tend to be very specific to particular soil types and parent materials (PMs), so cannot be applied universally and due consideration is required of the uncertainties of bioaccessibility models based on small numbers of analyses.

The British Geological Survey (BGS) holds extensive geochemical data and archives of soils collected under the Geochemical Baselines (G-BASE) Project (Johnson et al., 2005). If MLR or linear regression (LR) models could be set up using the soil geochemical data as predictor variables and As bioaccessibility data from selected *in vitro* bioaccessibility testing of soils, it would then, theoretically, be possible to predict the As bioaccessibility at a national scale and produce As bioaccessibility maps.

In this feasibility study, predictive regression modelling between bioaccessible As and a range of total elemental compositions and soil properties was executed for (1) the anthropogenically contaminated London (Scheib et al., 2011), Swansea (Marchant et al., 2011) and Glasgow (Broadway et al., 2010; Farmer et al., 2011) urban areas and (2) geogenic As associated with middle Jurassic ironstones and associated soil parent materials in the English East Midlands (Wragg et al., 2007) and Cretaceous ironstones in eastern England (Breward, 2007) and their associated soil parent materials (PMs). The objective was to assess the potential for developing bioaccessible As datasets derived from the BGS National and Urban Soil Chemistry datasets (Appleton et al., 2008; BGS, 2009; Appleton, 2011; Appleton and Adlam, 2012).

2. Geogenic and anthropogenic sources of arsenic in England and Wales

In central England concentrations above the residential As SGV (32 mg kg^{-1}) stretch from the Humber Estuary in the north all the way to the Bristol Channel in the southwest (Figure 1; Appleton et al., 2008; Rawlins et al., 2012). The majority relate to early Jurassic calcareous ironstones of the Northampton Sand and Marlstone Rock formations (Wragg et al., 2007). The most northern part of the feature relates to the Lower Cretaceous Claxby and Lower

Jurassic Frodingham ironstones (Palumbo-Roe et al., 2005; Breward, 2007) which occur in predominantly mudstone sequences. In sedimentary rocks, As is usually associated with iron and clay minerals, sulphides, organic matter and phosphates and as a consequence tends to be enriched in sedimentary ironstones, mudstones and coals. The association with phosphates is due to the chemical similarity between the PO_4^{3-} and AsO_4^{3-} anionic groups, while a strong association with sedimentary iron oxides is mainly due to the very low solubility of FeAsO_4 and also strong affinity for Fe oxyhydroxides (Cornell and Schwertmann, 2003)

Anthropogenic sources of As include coal combustion, sulphide ore roasting and smelting, pig and poultry sewage and some phosphate fertilisers (Reimann and de Caritat, 1998). The largest cluster of elevated total As concentrations can be seen throughout the southwest of England, which are related to Cu-Sn mineralisation associated with granite intrusions as well as the legacy of mining and mineral processing in the area, including arsenic refining (Klinck et al., 2005; Palumbo-Roe and Klinck, 2007). For the majority of urban areas, As concentrations are low and below the SGV (Flight and Scheib, 2011), although higher total As concentrations have been recorded in the urban topsoils from Swansea, Northampton, Scunthorpe, Corby, Sheffield, Manchester and Hull (Flight and Scheib, 2011; Scheib et al., 2011). The concentration of total As in Swansea topsoils is above the SGV for almost the entire conurbation, largely due to the city's industrial legacy of non-ferrous smelters processing copper (Cu), As, lead (Pb), zinc (Zn), silver (Ag) and tin (Sn) (Marchant et al., 2011). Mining of Jurassic ironstones occurred within and close to the Corby, Northampton, Scunthorpe and Wellingborough urban centres, but only Corby and Scunthorpe had major iron foundries.

3. Materials and methods

3.1 Sample selection

In order for the predictive regression models to be robust, it is necessary to ensure that the samples used for bioaccessibility testing are representative of the region under study.

The forty nine samples from London were selected to cover a range of total As and Pb concentrations, but data for three samples were not used in this study because the extremely high Pb concentrations meant that the As determinations were not reliable due to an XRF analytical interference. Twenty five topsoils from the Swansea urban area (Morley and Ferguson, 2001; Marchant et al., 2011) were selected to represent the wide range of total As and As/Fe ratios. Data for regression modelling in the Glasgow area is for the G-BASE samples reported in Broadway (2008), Broadway et al. (2010) and Farmer et al. (2011) which were originally selected to provide a range of total Cr, Pb or As concentrations. Only nine of the 21 bioaccessible As concentrations reported by Broadway in his PhD thesis (Broadway, 2008) were above the limit of detection of the ICP-AES analytical method (5 mg kg^{-1}).

The Northampton urban area is underlain by the Middle Jurassic ironstones of the Northampton Sand Formation (NSF) and associated Upper Lias Clays and Great Oolite Group limestones. The soil geochemical data set, containing 276 samples and consisting of 43 major and trace elements, pH, organic carbon and available phosphorus (Olsen method) was subjected to hierarchical clustering. The data were mean centred and scaled with Euclidean distance linkage using Ward's method. The resulting dendrograms suggested the existence of 4 distinct clusters in the Northampton data. A total of forty nine were chosen for further preparation and bioaccessibility testing. Twenty four topsoils with a range of relatively high total As were selected to represent the Cretaceous Claxby ironstone and associated ferruginous clays and sandstones in Lincolnshire.

3.2 Sample collection, preparation and determination of total concentrations

Topsoil samples from the Glasgow, London, Northampton and Swansea urban areas were collected from open ground on a 500 m grid at a density of approximately 4 samples per km²; samples for the Lincolnshire (Claxby) rural area in eastern England were collected at a density of approximately 1 sample per km². At each site, composite samples, based on 5 sub-samples taken at the centre and four corners of a 20 m square, were collected from the 5–20 cm depth. Approximately 40 chemical elements were determined in the <2 mm size fraction of the topsoils. Sample preparation, X-ray fluorescence analytical methods, and quality control procedures are described in Allen et al. (2011) and Johnson (2011).

3.3 Bioaccessible arsenic

Once selected for bioaccessibility testing, the samples were further sieved to <250 µm as this particle size fraction is considered to be the optimum size to adhere to children's hands (Duggan et al., 1985). The <250 µm fraction of the samples was assessed for As bioaccessibility As (B-As) using the Unified BARGE Method (UBM), an *in vitro* physiological gastro-intestinal (GI) simulation previously described by Wragg et al. (2011) and Denys et al. (2012). The UBM methodology has been validated against a Juvenile swine model for As, cadmium (Cd) and Pb (Denys et al., 2012) in a study of 16 different soils contaminated by mining and smelting practices, including the reference material NIST 2710. Correlation between the relative bioavailability and bioaccessibility of As was highly significant, both for the gastric and the gastro-intestinal phases, the slopes of the regression were not significantly different from 1 and the intercepts of the regression were not significantly different from zero.

Minor operational differences in the method occurred between the extraction of the Northampton and Glasgow samples and those from Swansea, Claxby and London because

the extractions were carried out at different phases of development of the UBM. The differences are as follows:

- The initial stomach phase pH for the Swansea, Claxby and London samples was fixed at 1.2 (according to Denys et al, 2012), whereas for the Northampton and Glasgow samples the tolerance on the initial stomach phase pH was 1.2 – 1.4 (according to Wragg et al, 2011); and
- The speed and length of time which the samples were spun at during the separation stage (4500 g for 15 minutes for the Swansea, Claxby and London samples compared to 3000 g for 5 minutes for the Northampton and Glasgow samples).

The UBM methodology provides samples for analysis from both the ‘stomach’ (BS) and ‘stomach & intestine’ (BSI) phases, equating to gastric and intestine compartments, although absorption of available As occurs in the small intestine of the GI tract. For the purposes of this study the highest concentration found in either of the two extracts (assigned as B-As in this study) was used for data interpretation; this was usually the BSI value.

Analysis of bioaccessibility extracts for the London, Northampton, Swansea and Claxby samples was carried out as described by Wragg et al. (2011) using a Thermo Elemental ExCell quadrupole ICP-MS instrument in combination with a Cetac ASX-510 autosampler, according to the operating conditions previously described by Watts et al. (2008). The Glasgow samples were analysed by ICP-AES Varian/Vista AX CCD simultaneous instrument with dedicated Varian SPS-5 Auto-sampler according to the operating conditions previously described (Wragg, 2005).

The quality control (QC) of the bioaccessibility extractions was monitored by carrying out replicate analyses of the BGS guidance soil BGS 102 (Wragg, 2009). At present there are no certified reference materials for bioaccessible As, however, the soils used for QC checks in this study have been the subject of an international inter-laboratory trial (Wragg et al., 2009;

2011), which has generated the reference values used in this study. Within every batch, of a maximum of 10 unknown samples, one duplicate, one quality control soil and one blank were extracted.

For all sample batches, London (n = 8), Glasgow (n = 4), Northampton (n=7), Swansea and Claxby batch (n = 8) replicate bioaccessibility values for BGS 102 were in good agreement with the consensus values of $4.5 \pm 1.3 \text{ mg kg}^{-1}$ and $5.4 \pm 2.4 \text{ mg kg}^{-1}$ for the 'stomach' and 'stomach & intestine' compartments respectively (Wragg et al., 2009; 2011). Bioaccessibility values obtained for the two compartments for the London sample batch were $3.8 \pm 0.1 \text{ mg kg}^{-1}$ and $3.3 \pm 0.3 \text{ mg kg}^{-1}$ respectively; $4.8 \pm 0.3 \text{ mg kg}^{-1}$ and $4.3 \pm 0.3 \text{ mg kg}^{-1}$ for Glasgow; for Northampton 4.92 ± 0.24 and $3.24 \pm 0.28 \text{ mg kg}^{-1}$ and for the batch which included Swansea and Claxby 4.51 ± 0.55 and $3.43 \pm 0.21 \text{ mg kg}^{-1}$ respectively.

For the soils collected from London, extraction duplicates were carried out for both the 'stomach' (n=13) and the 'stomach & intestine' (n=6) phases of the UBM. The mean As repeatability in the 'stomach' and the 'stomach & intestine' phases was 5.0% and 5.4%. Mean repeatability for the duplicate soils (n=21) extracted from the Northampton, Swansea and Claxby sample batch were 2.5 % in the 'stomach' and 2.3 % in the 'stomach & intestine' compartments of the UBM.

The UBM method detection limits for As were calculated as five times the average blank measurement. The calculated detection limits for the two extraction phases for As in the Northampton, Swansea, Claxby sample batch, equate to 0.5 mg kg^{-1} in the 'stomach' compartment and 1.5 mg kg^{-1} in the 'stomach & intestine' compartment of the UBM. For the London sample batch, the calculated detection limits for the two extraction phases for As equate to 0.7 and 0.95 mg kg^{-1} in the 'stomach' and the 'stomach & intestine' phases respectively. For the Glasgow sampling area the calculated detection limits were calculated

as 5.8 mg kg^{-1} and 5.2 mg kg^{-1} for the two compartments. All blank UBM extractions returned values below the method detection limits.

Differences between the As UBM bioaccessibility values before and after the stomach phase pH tolerance was fixed to 1.2 ± 0.05 were checked by comparing the results for the quality control check sample, BGS102. For both the 'stomach' and 'stomach & intestine' compartments the method with wider pH tolerance (Glasgow and Northampton) gave As bioaccessibility values that were in the normal range of the BGS 102 guidance soil and there were no consistent biases between the two sets of conditions.

3.4 Comparison of UBM and BGS-PBET bioaccessibility data

Topsoil bioaccessible As (B-As) determined by the UBM method for (i) geogenic ironstone and associated PMs and (ii) anthropogenically impacted urban areas (London, Glasgow, Swansea) are compared with bioaccessible As data (PB-As) determined using the BGS-PBET method (Cave et al., 2003) for (a) the Northampton Sand Formation (NSF) ironstones (PBET-NSF) from Wellingborough (Wragg et al., 2007) and rural Northamptonshire (Palumbo-Roe et al., 2005), (b) Tamar rural soils and Devon Gt. Consols Mine area medians (PBET-TAMA) (Klinck et al., 2005; Palumbo-Roe and Klinck, 2007), (c) the Lower Jurassic Frodingham ironstones in the Scunthorpe urban area (PBET-SCUN) (O'Donnell, 2005; Palumbo-Roe et al., 2005) and (d) the Cardiff urban area (Cave et al., 2003).

3.5 Statistical analysis

LR and MLR analysis in MINITAB® was used to model the relationship between bioaccessible As (B-As, <250 micron fraction) with total As, Al, Fe, K, Mg, Mn, Na, P, Si, Ti in the <2mm fraction of topsoil samples. pH was included as a variable for London and

Northampton, and pH and TOC for Glasgow only; pH and TOC data were not available for the other datasets.

Regression analysis in this study was carried out in the following stages:

- i. Use Pearson correlation coefficients between dependent variable (B-As) and potential independent variables (As, major elements, and pH, TOC and Olsen-P when data is available) to select those independent variables with the strongest linear relationship with B- As.
- ii. For each data set, use the exploratory MLR procedures of stepwise linear regression and best subset regression to identify the most important independent variables, bearing in mind that these procedures have limitations when the sample size (n) is < 5 times the number of independent variables. Repeat the MLR analysis after removing independent variables from the model if the b-coefficients were not significant and/or the ANOVA p-value for an independent variable was >0.05. The objective was to generate models which are geochemically and mineralogically justifiable. Residuals plots were used to assess whether residuals were approximately normally distributed.

Modelling bioaccessible As from total element concentrations in topsoils requires the application of regression models between B-As in the <250 μm fraction with total As and other chemical variables in the <2 mm fraction of topsoils. The relationship between total element concentrations in the <2mm and <250 μm fractions was investigated using data for London and Swansea; the results are reported in Appendix A (Supplementary data).

For Figure 3 where data populations are distinguished based on a plot of bioaccessible As against total As, the boundaries for the data populations were drawn using a two dimensional density estimate (Bowman and Azzalini, 1997) as implemented in the R programming language (R Development Core Team, 2011) in the programme package “sm”.

4. Results and discussion

4.1 Summary statistics

The mean, median and ranges of total As, bioaccessible As (B-As) and the bioaccessible fraction (%As-BAF) are given in Table 1. Median soil pH is 5.4, 6.8 and 6.4 for Glasgow, London and Northampton, respectively; no pH data is available for the Swansea or Claxby soil samples. The proportion of As that is bioaccessible (%As-BAF in Table 1) is higher (median 19-28%) in the Glasgow, London and Swansea urban areas where it is likely to be derived predominantly from anthropogenic sources, than in the Northampton and Lincolnshire (Claxby) areas (median 5-9%) where the As is likely to be derived mainly from Jurassic and Cretaceous ironstones and associated PMs.

4.2 Regression analysis

The LR modelling procedures described above were followed leading to the production of a range of statistically and geochemically appropriate LR and MLR models for bioaccessible As (Tables 2 and 3). R^2 indicates the proportion of the variance in B-As accounted for by each regression model. All regression models are statistically significant ($p < 0.005$). In least squares linear regression, the method makes the assumption that all the uncertainty is associated with the y axis. As a consequence, high intercept values may reflect the magnitudes of the errors on the input parameters. Positive intercept values are theoretically impossible so in such cases it may be more appropriate to use models without fitted intercepts.

4.2.1 Glasgow

Stepwise regression and MLR using total As and the major elements as potential predictors indicate that only total As is significant with total As explaining 69% of the total variance.

The fitted LR model for bioaccessible As has a moderate intercept (3.82) and the As coefficients of the fitted and non-fitted models are very similar (0.165 and 0.218, respectively; Table 2). The twelve Glasgow samples with B-As concentrations below the limit of determination were not used in the regression analysis.

4.2.2 London

Stepwise regression using total As and the major elements, pH and LoI as potential predictors indicate that only total As, Fe₂O₃, LoI and pH are significant. MLR confirms these as statistically significant predictors with total As explaining 75% of the variance followed by pH (6%), Fe₂O₃ (3%) and LoI (3%). Whilst B-As and total As correlate strongly with Fe, the negative coefficient for Fe in the MLR model (Table 2) implies that Fe restricts bioaccessibility of As. Previous studies (Palumbo et al., 2005; Wragg et al., 2007) have shown that As associated with aged crystalline forms of Fe oxide has low mobility in ingestion bioaccessibility tests. Using total As as the sole predictor, the LR fitted (Figure A-1 and Table A-1 in Supplementary data) and non-fitted models for bioaccessible As in the London urban area (Table 2) have almost identical total As coefficients.

4.2.3 Swansea

Stepwise regression and MLR using total As and the major elements as potential predictors indicate that only total As is significant. Using total As as the sole predictor, the LR fitted (Figure A-2 and Table A-1 in Supplementary data) and non-fitted models for bioaccessible As in the Swansea urban area (Table 2) have similar As coefficients, but the intercept (18.9 mg kg⁻¹ B-As) is impossibly high, probably influenced by the small number of samples with very high total As concentrations. For the subset with <600 mg kg⁻¹ As (Table 2), the intercept is less (7 mg kg⁻¹ B-As) but still positive with a magnitude similar to the uncertainties of the B-

As determinations (Wragg et al., 2011). Models without fixed intercepts would probably be more appropriate for this dataset.

4.2.4 Glasgow-London-Swansea combined

Arsenic coefficients for the LR models for the three urban areas where it is likely that most of the As is derived from anthropogenic sources are broadly the same. MLR and stepwise regression of the combined datasets (n = 80) confirms that As is the only significant predictor of B-As. Similar total As coefficients (0.165 and 0.218) and relatively high intercepts (7.0 and 2.6 mg kg⁻¹) characterise the models based on all the samples (n = 80) and on those samples with <600 mg kg⁻¹ total As (n = 78; Table 2). In both cases, the models with fitted intercepts explain a very high proportion of the total variance (91-92%, Table 2).

4.2.5 Northampton

The Northampton urban area is underlain by the ironstones of the Northampton Sand Formation, Upper Lias mudstones and calcareous rocks and limestones of the Great Oolite Group (Appleton and Adlam, 2012). Stepwise regression and MLR using total As and the major elements, pH and LoI as potential predictors indicate that only total As, P and pH are significant with total As explaining 53% of the variance followed by pH (14%), and P (5%). Using total As as the sole predictor, the LR fitted (Figure A-3 and Table A-1 in Supplementary data) and non-fitted models for bioaccessible As in the Northampton urban area (Table 3) have similar As coefficients. Using the subset of soil samples derived from the ironstones of the Northampton Sand Formation (NSF, n = 23), stepwise regression based on total As and the major elements, pH and LoI as potential predictors indicates that only P and pH are significant. This is confirmed by MLR which shows that P explains 36% of the variance and pH 32%. However, MLR models incorporating pH as a significant predictor

cannot be applied to the national BGS soil data because pH values have not been determined for all the BGS G-BASE soil samples.

4.2.6 Claxby ironstone and associated parent materials

This group of soils are derived from the Cretaceous Claxby ironstone, iron-rich mudstones and sandstones, together with a few Fe and As enriched samples derived from the Cretaceous Lower Greensand. Stepwise regression and MLR indicated that there are no statistically significant predictors of B-As.

4.2.7 Combined Northampton and Claxby ironstones and associated parent materials

Stepwise regression and MLR of the combined soil data for the Middle Jurassic NSF and Cretaceous Claxby ironstones and associated PMs (n = 73, Table 3) produced a model with total As, Fe (negative coefficient), Mg and P explaining 41%, 4%, 0.3% and 18%, respectively of the variance of B-As. The LR model based on total As as the sole predictor has a very low As coefficient (0.04, Table 3). Combining just the soil samples derived from the NSF and Claxby ironstones (n = 42), produced an MLR model in which total As, Ca, Fe and P explained 40%, 17%, 13% and 5% respectively of the total variance of B-As. The LR model with total As as the sole predictor again has a very low coefficient (0.03), which is compatible with the low %B-As (Table 1) for soils derived from these ironstones and associated PMs.

4.3 Comparison of UBM bioaccessibility for soils enriched in arsenic related to geogenic and anthropogenic processes

The summary percentage bioaccessibility statistics (Table 1), LR models (Tables 2 and 3) and Figure 2 clearly illustrate that soil samples enriched in As as a result of anthropogenic

processes (metal processing, disposal of coal ash, and background urban contamination) form a population that is significantly different to that for soils enriched in As principally as a result of geogenic controls. Soil samples from Glasgow and London, which probably reflect low level As enrichment resulting from disposal of coal ash and background urban contamination, fall within the same area of Figure 2, whilst the samples for Swansea follow the same trend but at much higher total As and B-As concentrations, reflecting the widespread impact of metal processing in the Swansea area. In contrast, the soil samples from Northampton and those from Claxby ironstone and associated PMs plot in a completely different field on the LnB-As vs. LnAs plot (Figure 2). There is overlap between the two groups of soils enriched in As as a result of geogenic processes, with soils derived from Claxby ironstones and associated PMs having higher total As (and Fe) and B-As but following the same broad trend as the Northampton soils derived from the Northampton Sand Formation ironstones and associated PMs (Figure 2).

4.4 Comparison of UBM data with published BGS-PBET bioaccessibility data

The mean, median and ranges of total As, BGS-PBET bioaccessible As (PB-As) and the bioaccessible fraction (%As-PB-BAF) are given in Table 4. The BGS-PBET data for the Northampton Sand Formation ironstones from Wellingborough and rural Northamptonshire generally fall within the same field as the UBM data for the NSF samples from Northampton and the Cretaceous ironstones and associated PMs, although some of the BGS-PBET data indicate lower bioaccessibility than the UBM bioaccessible As data (Figure 3). The BGS-PBET data for the Frodingham ironstones in the Scunthorpe area mainly fall within the UBM ‘geogenic’ (ironstone and associated parent materials) field although a few have higher relative bioaccessibility and plot within the anthropogenic field. These soil samples are located close to the iron foundries in Scunthorpe so may be impacted by diffuse pollution

derived from metal processing and consequently have higher As bioaccessibility, as observed in the Swansea area. BGS-PBET data for regional soils from the Tamar catchment generally fall within the 'geogenic' field and its extension towards higher total As and B-As concentrations (Figure 3). However, there are a group of 5 samples from the immediate vicinity of the Devon Great Consols Mine with greater than 500 mg kg⁻¹ As which overlap with the UBM 'anthropogenic' field and have higher bioaccessibility of 15-20% compared with an average of about 7% for 'background' soils from the Tamar catchment (Figure 3). Arsenic was produced at Devon Great Consols arsenic refinery between 1865 and 1930 by burning arsenopyrite, so it is likely that soils in the immediate vicinity will be impacted by As oxide soot, which is likely to have a higher bioaccessibility than arsenopyrite and its natural weathering products. BGS-PBET data for three soils collected from fields to the north of Cardiff plot within the UBM 'geogenic' field and have an average bioaccessibility (%As-PB-BAF) of 13%. One sample from woodland outside the urban area has a much high PB-As (11.6 mg kg⁻¹) and %As-PB-BAF (30%) although the soil has a very low pH (3.2), which may explain the higher bioaccessibility. Fourteen samples from the Cardiff urban area collected from industrial sites, residential gardens and allotments, plot within the UBM 'anthropogenic' field (Figure 3) and have an average bioaccessibility (%As-PB-BAF) of 25%. Only one urban soil sample, collected from reclaimed land near the Cardiff East Moors steel works (Brown, 2001) has an unusually low bioaccessibility (5% As-PB-BAF).

4.5 Comparison of the As bioaccessibility data with other studies

To date this study is the most comprehensive survey of the bioaccessibility of As in the UK in terms of the number of soil samples analysed, the regional coverage and the use of a validated bioaccessibility test. Comparing the results from this work to other studies is not straight forward because of the different methods used for measuring bioaccessibility;

however, making the assumption that there is broad agreement between methods some generalised comparisons can be made. Cave and Wragg (2003) carried out a preliminary assessment of the bioaccessibility of As in UK soils using a modified version of Ruby's method (Cave, 2003). In a similar manner to this study, they found that the As from areas with greater anthropogenic disturbance (Cardiff city and the As mine at Devon Great Consols in Devon) the bioaccessibility was generally higher (median values 10-22%) than in Wellingborough where the As in the soil was naturally derived from underlying ironstones (median value 6%). In other countries most of the As bioaccessibility studies have concentrated on anthropogenic contamination. In the US a recent study (EPA, 2010) has measured the in-vivo bioavailability (using juvenile swine) of As in 29 soils from 11 contaminated sites mostly associated with mining activity which gave a relative bioavailabilities ranging from 10-60%. In Australia (Juhasz et al., 2007b) measured the bioaccessibility of As in soils contaminated with As based herbicides and cattle dips, soils from former gold mine sites and from highly mineralised sites containing geogenic As using the glycerine based bioaccessibility method (Drexler, 2007). The bioaccessibility of As within each soil type was variable but overall all of the soils had relative bioaccessible As <50%.

4.6 Production of bioaccessibility maps using LR models

The combined Glasgow-London-Swansea (GLS) model ($B\text{-As} = 0.175 \text{ total As}$) was used to produce an estimated B-As map from an estimated topsoil total As dataset for the Swansea urban area created by interpolating log transformed total As data to a 100m grid using an IDW algorithm (Appleton, 2011). The resultant bioaccessible As map (Figure 4) shows that an extensive part of central Swansea is characterised by estimated B-As concentrations greater than 10 mg kg^{-1} but only a relatively restricted area has B-As exceeding 30 mg kg^{-1} .

5. Conclusions

The relationship between bioaccessible and total As in UK urban areas, such as London and Glasgow, where soil As concentrations are probably impacted by spreading of coal ash and other types of background anthropogenic contamination, appears to be consistent with data for urban centres, such as Swansea, which have a long history of non-ferrous metal processing including As refining.

Total As is the only statistically significant predictor in the Glasgow, London and Swansea urban areas explaining 91-92% of the B-As variance, whereas for ironstone domains models based on total As as the sole predictor explain only about 40% of the variance of B-As.

Comparison of the results of the new UBM bioaccessibility data with previously published BGS-PBET data appear to indicate that the results of the two methods are broadly similar for soils impacted by geogenic 'contamination'. The origin and mineralogy of the soils can impact on the bioaccessibility of As, so bioaccessible As concentrations estimated using these models should be treated with some caution. However, the broad compatibility of the London, Glasgow and Swansea regression models suggests that application of these models to estimate bioaccessible As in UK soils impacted by normal types of background urban contamination and non-ferrous metal processing should be relatively accurate. Equally, application of the 'geogenic' model based on bioaccessibility data for the Jurassic and Cretaceous ironstones to all sedimentary ironstone terrains in the UK appears to be justified.

Validation of the models reported in this study by comparable studies in other urban areas of the UK, such as Corby (Scheib and Nice, 2009), Sheffield (Rawlins et al., 2005) and Humberside (Rawlins et al., 2006) should be carried out along with re-analysis of some of the

Tamar and Wellingborough samples by the UBM method to assess in more detail the relationship between As bioaccessibility determined by these two methods.

Oxalate extractable As or near infra-red spectral data indicating the different forms of Fe oxides in the soil may be significant predictors of As bioaccessibility. Unfortunately this information is not available in the UK national G-BASE dataset.

The results of this study draw broad conclusions on the bioaccessibility of As in anthropogenic (urban) and geogenic environments in the UK. These results can be used as part of a lines of evidence approach to localised risk assessment but should in no way be used to replace bioaccessibility testing at individual sites where local conditions may vary considerably from the broad overview presented in this study.

Acknowledgements

This paper is published with the permission of the Director of the British Geological Survey (Natural Environment Research Council). The contribution of all BGS staff involved in the collection and analysis of samples from the G-BASE survey is gratefully acknowledged.

Andreas Scheib (BGS) tested some of the MLR models. The Glasgow data were obtained as part of the PhD research of A. Broadway, University of Edinburgh. Analyses were carried out by staff of the BGS Analytical Geochemistry Facility. Barbara Palumbo-Roe (BGS) and two anonymous reviewers are thanked for suggesting useful improvements to an earlier version of this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi: [insert link]

References

Allen MA, Cave MR, Chenery SRN, Gowing CJB, Reeder S. Sample Preparation and Inorganic Analysis for Urban Geochemical Survey Soil and Sediment Samples. In: Johnson C, Demetriades A, Locutura J, Ottesen RT, editors. Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, 2011, pp. 28-46.

Ander EL, Cave MR, Johnson CC, Palumbo-Roe B. Normal background concentrations of contaminants in the soils of England. Available data and data exploration. British Geological Survey, 2012.

Appleton JD. User Guide for the British Geological Survey Urban Soil Chemistry Dataset. British Geological Survey, Keyworth, Nottingham, UK, 2011.

Appleton JD, Adlam KAM. Geogenic control on soil chemistry in urban areas: A novel method for urban geochemical mapping using parent material classified data. Applied Geochemistry 2012; 27: 161-170.

Appleton JD, Rawlins BG, Thornton I. National-scale estimation of potentially harmful element ambient background concentrations in topsoil using parent material classified soil:stream-sediment relationships. Applied Geochemistry 2008; 23: 2596-2611.

Bacigalupo C, Hale B. Human health risks of Pb and As exposure via consumption of home garden vegetables and incidental soil and dust ingestion: A probabilistic screening tool. Science of the Total Environment 2012; 423: 27-38.

Basta NT, Foster JN, Dayton EA, Rodriguez RR, Casteel SW. The effect of dosing vehicle on arsenic bioaccessibility in smelter-contaminated soils. Journal of Environmental Science and Health Part A 2007; 42: 1275 - 1281.

BGS. National Soil Chemistry Dataset User Guide. British Geological Survey, Keyworth, Nottingham, UK, 2009.

Bowman AW, Azzalini A. Applied Smoothing Techniques for Data Analysis - The Kernel Approach with S-Plus Illustrations. Oxford: Clarendon Press, 1997.

Breward N. Arsenic and presumed resistate trace element geochemistry of the Lincolnshire (UK) sedimentary ironstones, as revealed by a regional geochemical survey using soil, water and stream sediment sampling. *Applied Geochemistry* 2007; 22: 1970-1993.

Broadway A. Development of methodologies for soil metal bioaccessibility and human health risk. School of GeoSciences. Edinburgh, 2008.

Broadway A, Cave MR, Wragg J, Fordyce FM, Bewley RJF, Graham MC, et al.

Determination of the bioaccessibility of chromium in Glasgow soil and the implications for human health risk assessment *Science of The Total Environment* 2010; 409: 267-277.

Brown SE. Geochemical baseline data for the urban area of Cardiff. British Geological Survey, Keyworth, Nottingham, UK, 2001.

Caboche J. Validation d'un test de mesure de bioaccessibilité. Application à quatre éléments traces métallique dans les sols: As, Cd, Pb et Sb. Science Agronomique. PhD. L'Institut National Polytechnique de Lorraine, Nancy, 2009, pp. 348.

Cave MR, Taylor H, Wragg J. Estimation of the Bioaccessible Arsenic Fraction in Soils using Near Infra Red Spectroscopy. *Journal of Environmental Science and Health Part A* 2007; 42: 1293-1301.

Cave MR, Wragg J, Denys S, Jondreville C, Feidt C. Oral Bioavailability. In: Swartjes F, editor. *Dealing with Contaminated Sites: From Theory towards Practical Application*. Springer, 2011.

Cave MR, Wragg J, Palumbo B, Klinck BA. Measurement of the Bioaccessibility of Arsenic in UK soils. Environment Agency, 2003.

Cornell RM, Schwertmann U. *The Iron Oxides - Structure Properties, Reactions, occurrences and Uses*. Weinheim: VCH Publishers, 1996.

DEFRA-EA (Department for the Environment Food and Rural Affairs and the Environment Agency). Contaminants in Soil: Collation of Toxicological Data and Intake Values for Humans. Arsenic. Department for the Environment, Food and Rural Affairs, 2002.

Denys S, Caboche J, Tack K, Rychen G, Wragg J, Cave M, et al. In Vivo Validation of the Unified BARGE Method to Assess the Bioaccessibility of Arsenic, Antimony, Cadmium, and Lead in Soils. *Environ Sci Technol* 2012; 46: 6252-60.

Duggan MJ, Inskip MJ, Rundle SA, Moorcroft JS. Lead in Playground Dust and on the Hands of Schoolchildren. *Science of the Total Environment* 1985; 44: 65-79.

EPA. Relative bioavailability of arsenic in soils at 11 hazardous waste sites using in vivo juvenile swine method. U.S. Environmental Protection Agency, Washington, USA, 2010, pp. 56.

Farmer JG, Broadway A, Cave MR, Wragg J, Fordyce FM, Graham MC, et al. A lead isotopic study of the human bioaccessibility of lead in urban soils from Glasgow, Scotland. *Science of the Total Environment* 2011; 409: 4958-4965.

Flight DMA, Scheib AJ. Soil Geochemical Baselines in UK Urban centres: The G-BASE Project. In: Johnson C, Demetriades A, Locutura J, Ottesen RT, editors. *Mapping the Chemical Environment of Urban Areas*. Wiley-Blackwell, Oxford, 2011, pp. 186-206.

Intawongse M, Dean JR. In-vitro testing for assessing oral bioaccessibility of trace metals in soil and food samples. *Trac-Trends In Analytical Chemistry* 2006; 25: 876-886.

Johnson CC. Understanding the Quality of Chemical Data from the Urban Environment – Part 1: Quality Control Procedures. In: Johnson C, Demetriades A, Locutura J, Ottesen RT, editors. *Mapping the Chemical Environment of Urban Areas*. Wiley-Blackwell, Oxford, 2011, pp. 61-76.

Johnson CC, Breward N, Ander EL, Ault L. G-BASE: Baseline geochemical mapping of Great Britain and Northern Ireland. *Geochemistry: Exploration, Environment, Analysis* 2005; 5: 347-357.

Juhasz AL, Smith E, Weber J, Naidu R, Rees M, Rofe A, et al. Effect of soil ageing on in vivo arsenic bioavailability in two dissimilar soils. *Chemosphere* 2008; 71: 2180-2186.

Juhasz AL, Smith E, Weber J, Rees M, Rofe A, Kuchel T, et al. Comparison of in vivo and in vitro methodologies for the assessment of arsenic bioavailability in contaminated soils. *Chemosphere* 2007a; 69: 961-966.

Juhasz AL, Smith E, Weber J, Rees M, Rofe A, Kuchel T, et al. In vitro assessment of arsenic bioaccessibility in contaminated (anthropogenic and geogenic) soils. *Chemosphere* 2007b; 69: 69-78.

Klinck B, Palumbo B, Cave MR, Wragg J. Arsenic dispersal and bioaccessibility in mine contaminated soils: a case study from an abandoned arsenic mine in Devon, UK. British Geological Survey, 2005.

Marchant BP, Tye AM, Rawlins BG. The assessment of point-source and diffuse soil metal pollution using robust geostatistical methods: a case study in Swansea (Wales, UK). *European Journal of Soil Science* 2011; 62: 346-358.

Martin I, De Burca R, Morgan H. Soil Guideline Values for inorganic arsenic in soil. Environment Agency of England and Wales, 2009.

Meunier L, Koch I, Reimer KJ. Effect of particle size on arsenic bioaccessibility in gold mine tailings of Nova Scotia. *Science of the Total Environment* 2011a; 409: 2233-2243.

Meunier L, Koch I, Reimer KJ. Effects of dissolution kinetics on bioaccessible arsenic from tailings and soils. *Chemosphere* 2011b; 84: 1378-1385.

Mingot J, De Miguel E, Chacon E. Assessment of oral bioaccessibility of arsenic in playground soil in Madrid (Spain): A three-method comparison and implications for risk assessment. *Chemosphere* 2011; 84: 1386-1391.

Morley SE, Ferguson AJ. Geochemical baseline data for the urban area of Swansea. British Geological Survey, Keyworth, Nottingham, UK, 2001.

Nathanail CP. Professional Practice Note: Reviewing human health risk assessment reports invoking contaminant oral bioavailability measurements or estimates. . Chartered Institute for Environment and Health, 2009.

O'Donnell KE. Geochemical baseline data for the urban area of Scunthorpe. British Geological Survey, Keyworth, Nottingham, UK, 2005.

Oomen AG, Hack A, Minekus M, Zeijdner E, Cornelis C, Schoeters G, et al. Comparison of five in vitro digestion models to study the bioaccessibility of soil contaminants. *Environmental Science & Technology* 2002; 36: 3326-3334.

Palumbo-Roe B, Cave MR, Klinck BA, Wragg J, Taylor H, O'Donnell K, et al. Bioaccessibility of arsenic in soils developed over Jurassic ironstones in eastern England. *Environmental Geochemistry and Health* 2005; 27: 121-130.

Palumbo-Roe B, Klinck B. Bioaccessibility of arsenic in mine waste-contaminated soils: A case study from an abandoned arsenic mine in SW England (UK). *Journal of Environmental Science and Health Part A* 2007; 42: 1251 - 1261.

Paustenbach DJ. The practice of exposure assessment: A state-of-the-art review (Reprinted from *Principles and Methods of Toxicology*, 4th edition, 2001). *Journal of Toxicology and Environmental Health-Part B-Critical Reviews* 2000; 3: 179-291.

Pelfrene A, Waterlot C, Mazzuca M, Nisse C, Bidar G, Francis D. Assessing Cd, Pb, Zn human bioaccessibility in smelter contaminated agricultural topsoils (northern France). *Environmental Geochemistry and Health* 2011.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., 2011.

Rawlins BG, Lark RM, O'Donnell KE, Tye AM, Lister TR. The assessment of point and diffuse metal pollution of soils from an urban geochemical survey of Sheffield, England. *Soil Use And Management* 2005; 21: 353-362.

Rawlins BG, Lark RM, Webster R, O'Donnell KE. The use of soil survey data to determine the magnitude and extent of historic metal deposition related to atmospheric smelter emissions across Humberside, UK. *Environmental Pollution* 2006; 143: 416-426.

Rawlins BG, McGrath SP, Scheib AJ, Breward N, Cave M, Lister TR, et al. The advanced soil geochemical atlas of England and Wales. British Geological Survey, Keyworth, Nottingham, UK, 2012.

Reimann C, Caritat P. Chemical Elements in the Environment. Factsheets for the Geochemist and Environmental Scientist. Berlin: Springer-Verlag, 1998.

Reimann C, Filzmoser P, Garrett RG, Dutter R. Statistical Data Analysis Explained. Chichester: Applied Environmental Statistics with R. Wiley, 2008.

Rieuwerts JS, Searle P, Buck R. Bioaccessible arsenic in the home environment in southwest England. *Science Of The Total Environment* 2006; 371: 89-98.

Roussel H, Waterlot C, Pelfrene A, Pruvot C, Mazzuca M, Douay F. Cd, Pb and Zn oral bioaccessibility of urban soils contaminated in the past by atmospheric emissions from two lead and zinc smelters. *Arch Environ Contam Toxicol* 2010; 58: 945-54.

Sarkar D, Makris KC, Parra-Noonan MT, Datta R. Effect of soil properties on arsenic fractionation and bioaccessibility in cattle and sheep dipping vat sites. *Environment International* 2007; 33: 164-169.

Scheib AJ, Nice SE. Soil geochemical baseline data for the urban areas of Corby, Coventry, Derby, Leicester, Northampton, Nottingham and Peterborough in the East Midlands. British Geological Survey, Keyworth, Nottingham, UK, 2009.

Scheib C, Flight D, Lister TR, Scheib A, Breward N. London Earth: As in topsoils. G-BASE summary interpretation. British Geological Survey, Keyworth, Nottingham, UK, 2011.

Selinus O. Essentials of Medical Geology Impacts of the Natural Environment on Public Health: Elsevier Academic Press, 2005.

Tang X-Y, Zhu Y-G, Shan X-Q, McLaren R, Duan J. The ageing effect on the bioaccessibility and fractionation of arsenic in soils from China. *Chemosphere* 2007; 66: 1183-1190.

Watts MJ, Button M, Brewer TS, Jenkin GRT, Harrington CF. Quantitative arsenic speciation in two species of earthworms from a former mine site. *Journal of Environmental Monitoring* 2008; 10: 753 -759.

Wragg J. A Study of the Relationship between Arsenic Bioaccessibility and its Solid Phase Distribution in Wellingborough Soils. *Geography*. Nottingham, Nottingham, 2005.

Wragg J. BGS Guidance Material 102, Ironstone Soil, Certificate of Analysis. British Geological Survey 2009.

Wragg J, Cave M, Nathanail P. A Study of the relationship between arsenic bioaccessibility and its solid-phase distribution in soils from Wellingborough, UK. *Journal of Environmental Science and Health Part A* 2007; 42: 1303 - 1315.

Wragg J, Cave MR. In-vitro Methods for the Measurement of the Oral Bioaccessibility of Selected Metals and Metalloids in Soils: A Critical Review. Environment Agency, 2003.

Wragg J, Cave MR, Basta N, Brandon E, Casteel S, Denys Seb, et al. An Inter-laboratory Trial of the Unified BARGE Bioaccessibility Method for Arsenic, Cadmium and Lead in Soil. *Science of the Total Environment* 2011; 409: 4016-4030.

Wragg J, Cave MR, Taylor H, Basta N, Brandon E, Casteel S, et al. Inter-laboratory Trial of a Unified Bioaccessibility Procedure. British Geological Survey, 2009.

Yang JK, Barnett MO, Jardine PM, Basta NT, Casteel SW. Adsorption, sequestration, and bioaccessibility of As(V) in soils. *Environmental Science & Technology* 2002; 36: 4562-4569.

Table 1 Summary statistics for total As (<2 mm fraction), UBM bioaccessible arsenic (B-As, <250 μm fraction) and bioaccessible fraction (% As-BAF)

	N	Total As (mg kg^{-1})			B-As (mg kg^{-1})			%As-BAF		
		Range	Mean	Median	Range	Mean	Median	Range	Mean	Median
Glasgow	21	8-130	32	24	2.5-25.0	6.5	2.5	6-63	22	19
London	46	7-88	25	21	1.7-27.5	7.6	5.9	12-68	30	28
Swansea	25	55- 2161	313	170	13-313	66.6	42.0	10-44	25	22
Northampton (NSF and associated parent materials)	49	17-70	36	33	1.2-6.7	3.2	2.9	5-14	9	9
Claxby ironstone and associated parent materials	24	37-224	120	119	3.1-18.3	6.7	5.6	2-15	6	5

Table 2 Least squares linear regression models for bioaccessible arsenic: principally anthropogenic sources

Dataset	N	Intercept ¹	Model formula (<i>MLR models in italics</i>)	R ²
GLASGOW (G)	9	F	<i>B-As = 3.82 + 0.165 As</i>	69%
	9	NF	<i>B-As = 0.218 As</i>	90%
LONDON (L)	46	F	<i>B-As = - 13.8 + 0.350 As - 1.25 Fe₂O₃ + 0.480 LOI + 2.15 pH</i>	88%
	46	F	<i>B-As = -0.15 + 0.307 As</i>	75%
	46	NF	<i>B-As = 0.303 As</i>	91%
SWANSEA (S)	25	F	<i>BS As = 18.9 + 0.153 As</i>	91%
	25	NF	<i>BS As = 0.174 As</i>	93%
SWANSEA < 600 mg/kg As subset	23	F	<i>BS As = 7.28 + 0.203 As</i>	82%
	23	NF	<i>BS As = 0.228 As</i>	95%
G + L + S	80	F	<i>B-As = 7.04 + 0.165 As</i>	92%
	80	NF	<i>B-As = 0.175 As</i>	92%
G + L + S < 600 mg/kg As subset	78	F	<i>B-As = 2.576 + 0.2184 As</i>	91%
	78	NF	<i>B-As = 0.230 As</i>	94%

¹ F = fixed; NF = not fixed

Table 3 Least squares linear regression models for bioaccessible arsenic: principally geogenic sources

Dataset	N	Intercept¹	Model formula (<i>MLR models in italics</i>)	R²
Northampton	49	F	<i>BS-As = - 3.39 + 0.0359 As + 0.000507 P + 0.654 pH</i>	72%
	49	F	<i>BS As = 0.84 + 0.064 As</i>	53%
	49	NF	<i>BS As = 0.0847 As</i>	93%
NSF ironstone	23	F	<i>BS-As = - 5.54 + 0.000697 P + 1.20 pH</i>	69%
	23	F	<i>BS As = 0.6356 + 0.06795 As</i>	42%
	23	NF	<i>B-As = 0.0811 As</i>	93%
Northampton + Claxby	73		<i>BS As = - 1.20 + 0.0692 As - 0.000057 Fe + 0.000524 Mg + 0.00156 P</i>	66%
	73	F	<i>B-As = 2.02 + 0.0364 As</i>	41%
	73	NF	<i>B-As = 0.0561 As</i>	77%
NSF and Claxby ironstones	42	F	<i>BS As = 0.934 + 0.0642 As + 0.000079 Ca - 0.000048 Fe + 0.000728 P</i>	75%
	42	F	<i>BS As = 2.58 + 0.0280 As</i>	40%
	42	NF	<i>BS As = 0.0513 As</i>	80%

¹ F = fixed; NF = not fixed *

Table 4 Summary statistics for total As (<2 mm fraction), BGS-PBET bioaccessible arsenic (PB-As, <250 µm fraction) and bioaccessible fraction (% As-PB-BAF)

Area	N	Total As (mg kg ⁻¹)			PB-As (mg kg ⁻¹)			%As-PB-BAF		
		Range	Mean	Median	Range	Mean	Median	Range	Mean	Median
Cardigan (rural)	4	15-39	26	25	2.0–12	4.9	3.1	10–30	18	15
Cardigan (urban)	15	21-158	47	34	1.9–35	10.8	7.6	5–45	23	22
Devon Great Consols Mine	4	71-2105	1009	929	7–408	155	102	9–19	12	11
Tamar (rural)	94	9-865	94	32	0.7-126	8.7	2.5	3–27	9	9
Scunthorpe	40	20-179	42	34	2.4-17	5.2	4.3	3–30	15	14
Wellingborough	34	13-332	72	63	1.7-6	3.0	2.7	1–13	5	4

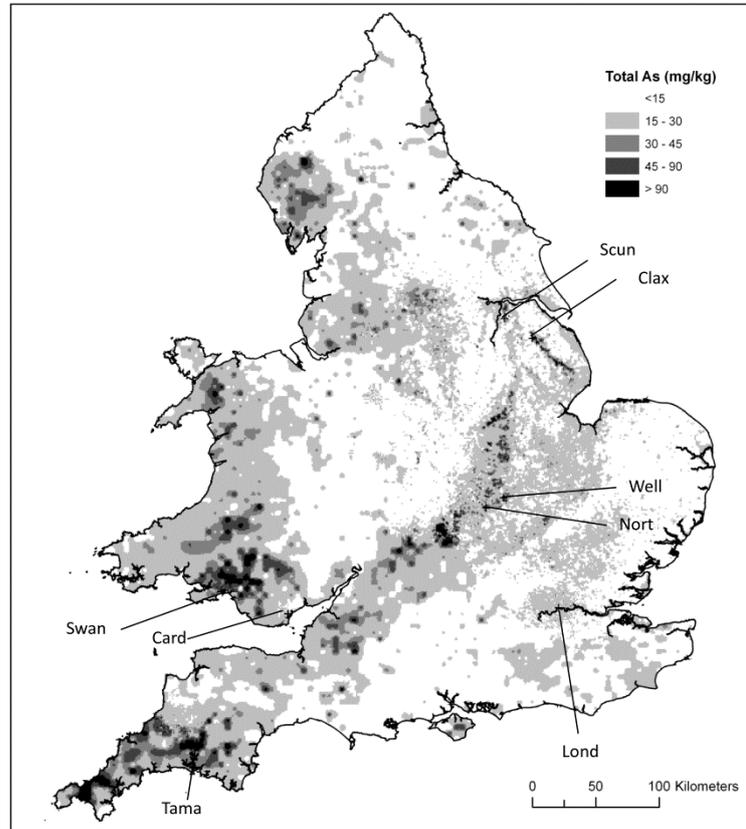


Figure 1 Estimated Total As in topsoils of England and Wales (IDW interpolation of LnAs BGS G-BASE and NSI topsoil data; Scun = Scunthorpe, Clax=Claxby, Well=Wellingborough, Nort=Northampton, Lond=London, Swan=Swansea, Card=Cardiff, Tama=Tamar)

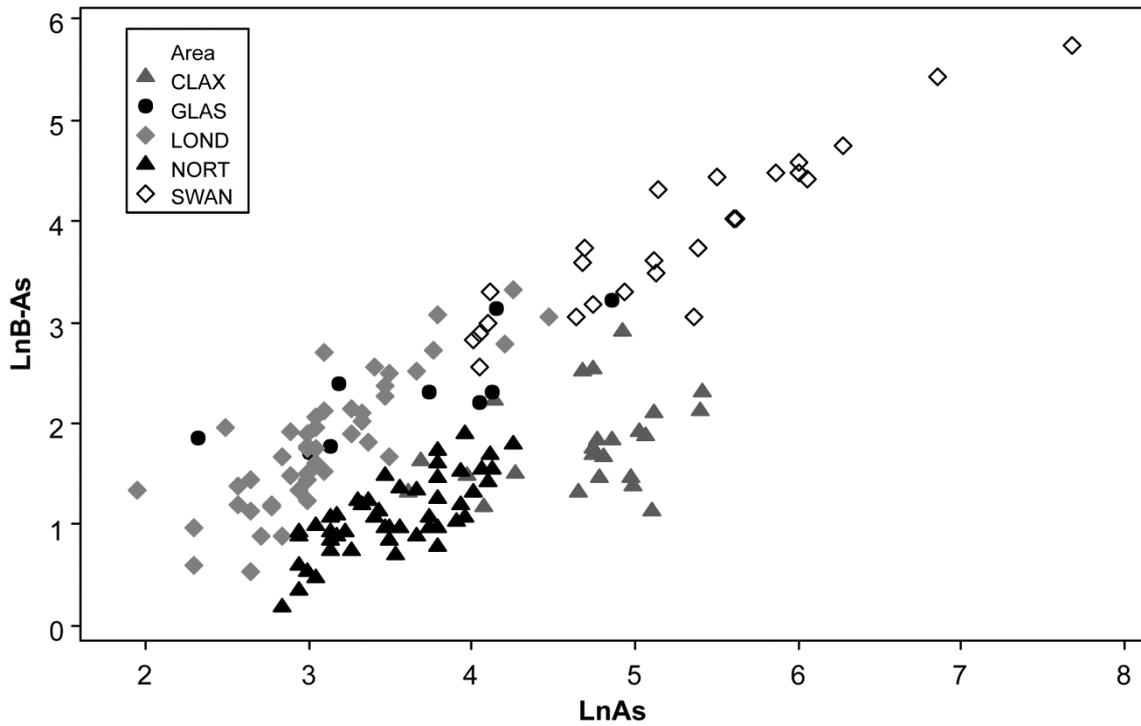


Figure 2 Relationship between Ln Total As (LnAs <2mm fraction) and Ln B-As (<250 micron fraction) for geogenic (ironstone and associated parent materials; CLAX=Claxby; NORT=Northampton) and anthropogenic (GLAS=Glasgow, LOND=London, SWAN=Swansea) domains

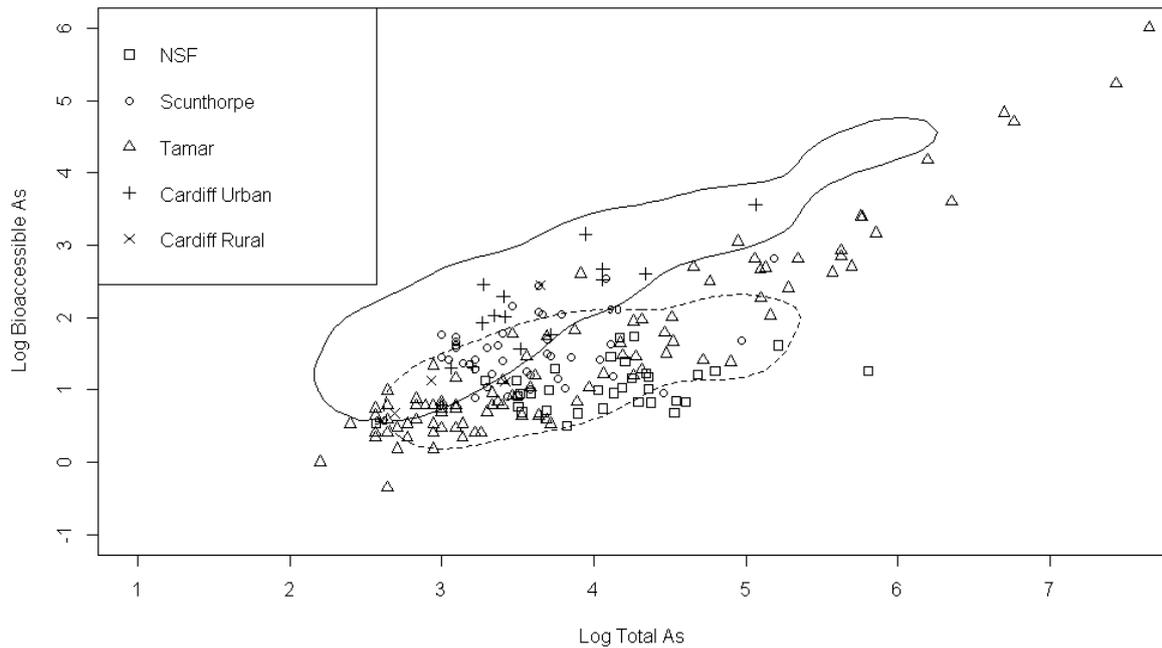


Figure 3 Relationship between total Ln Total As (<2mm fraction) and Ln B-As (<250 micron fraction) for UBM ironstone and associated parent materials domain (dashed line), UBM anthropogenic domain (London, Glasgow, Swansea; solid line), with BGS-PBET Northampton ironstone data from Wellingborough and rural Northamptonshire (NSF), Tamar rural soils and Devon Gt. Consols Mine area medians (Tamar), the Lower Jurassic Frodingham ironstones in the Scunthorpe urban area (Scunthorpe) and urban and rural topsoils from the Cardiff area.

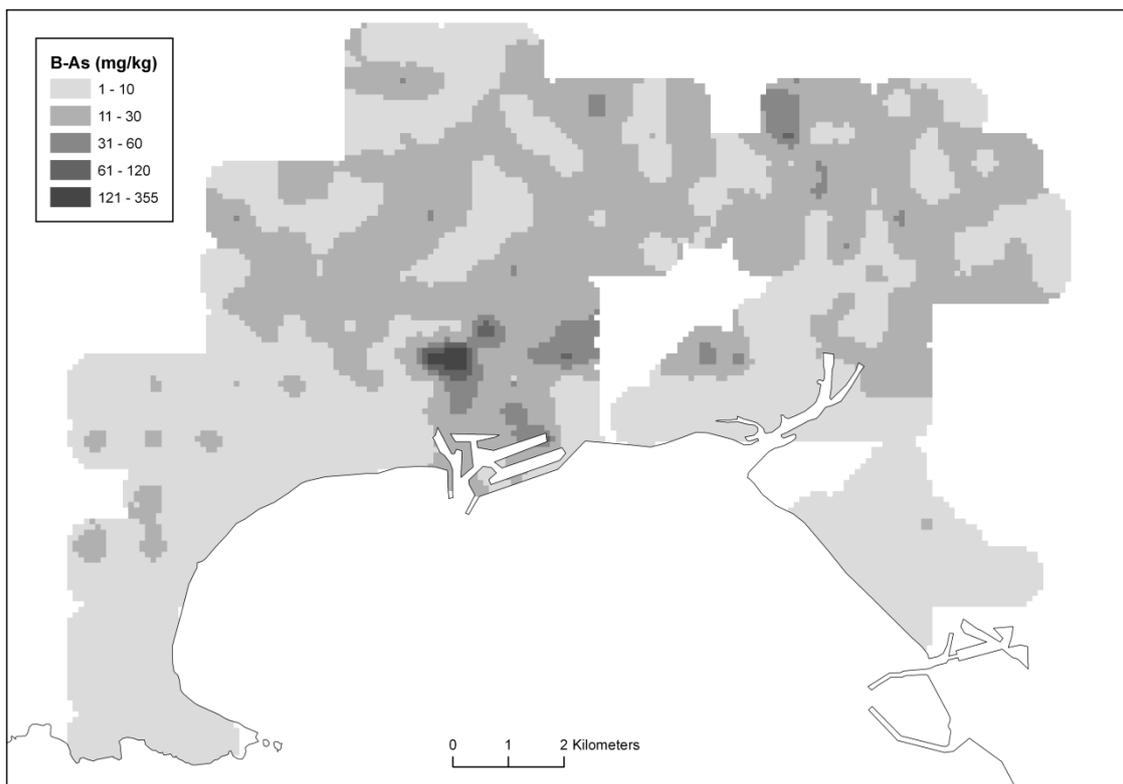


Figure 4 Estimated bioaccessible As in topsoils in the Swansea area (coastline: Ordnance Survey Strategi data © Crown copyright 2012)

Appendix A. Supplementary data

Section A-1

Linear regression models for London and Swansea anthropogenic domains and Northampton geogenic domain

Table A-1 Model statistics for linear regression models in Figures A-1, A-2, A-3

	n	Intercept Coeff.	Intercept SE	Intercept P	Slope Coeff.	Slope SE	Slope P	R2	adj.
Figure A-1	46	-0.157	0.791	0.843	0.306	0.027	<0.0005	75.1	74.5
Figure A-2	25	18.900	5.291	0.002	0.153	0.010	<0.0005	91.0	90.6
Figure A-3	49	0.844	0.340	0.017	0.064	0.009	<0.0005	53.3	52.3

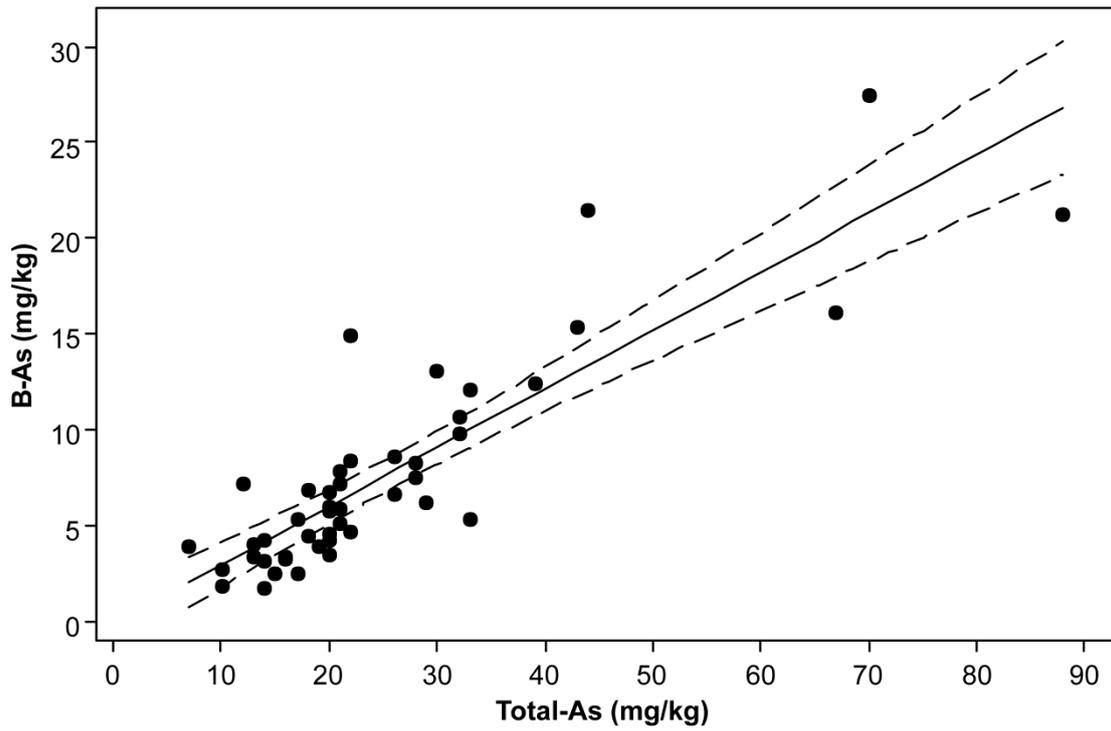


Figure A-1 Confidence limits (95%) for B-As (<250 micron fraction) vs. Total As (<2mm fraction) linear regression model for topsoils derived from London anthropogenic domain

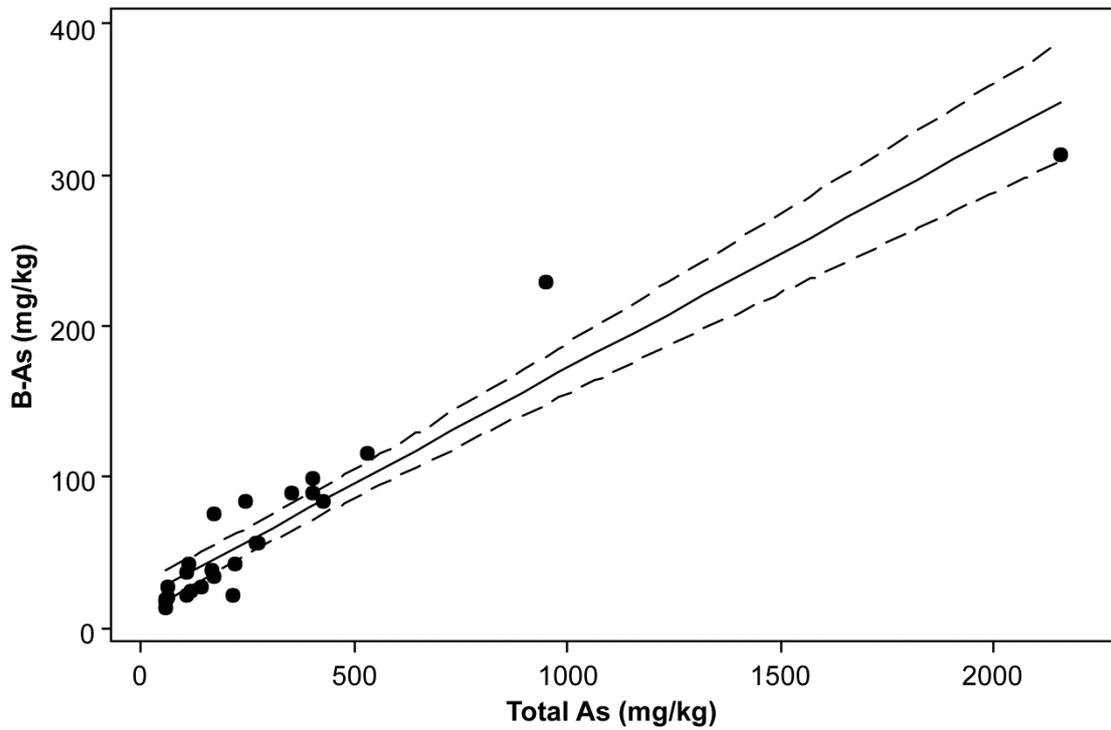


Figure A-2 Confidence limits (95%) for B-As (<250 micron fraction) vs. Total As (<2mm fraction) linear regression model for topsoils derived from Swansea anthropogenic domain

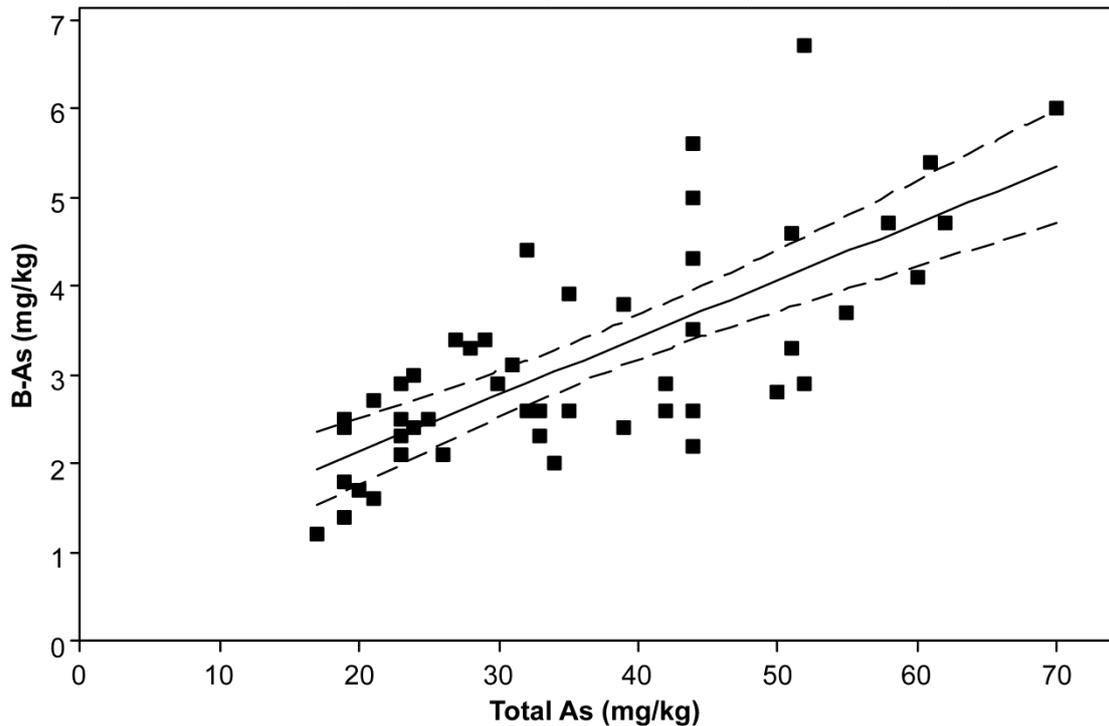


Figure A-3 Confidence limits (95%) for B-As (<250 micron fraction) vs. Total As (<2mm fraction) linear regression model for topsoils derived from Northampton geogenic domain

Section A-2

Comparison of models based on <2 mm and <250 μm fractions

Summary statistics for the <2mm and <250 μm fractions of the London samples (Table A-1) show that mean and median As are very similar in the <2mm and <250 micron fractions, whilst mean Si and Fe are about 18% higher in the <2mm fraction and mean Al and Mg slightly higher in the <250 μm fraction. The closed array (100%) effect will impact on these relationships (Reimann et al., 2008).

Mean and median As are similar in the <2mm and <250 μm fractions of the Swansea urban samples whilst Si is slightly higher in the <250 μm fraction and mean Al and Ca, are slightly higher in the <250 μm fraction (Table A-2).

The LR models for B-As vs As (<2mm) and total As (<250 μm) are very similar in both the London and Swansea areas (Figures A-4 and A-5). These results suggest that in the London and Swansea areas, size fraction is not an important factor for B-As vs. total As models. At high total As concentrations, grain size distribution may be a more important factor in some samples, if there is significant influence caused by the type of anthropogenic contamination or the types of soils developed on different PMs.

Table A-2 Summary statistics for <2mm and <250 µm fractions of topsoils from London urban area (n = 49)

Variable	Mean	Minimum	Median	Maximum
B-As <250 µm	7.5	1.7	5.8	27.5
As < 2mm	24	1	20	88
As <250 µm	25	1	19	100
Al ₂ O ₃ < 2mm	7.7	4.1	7.0	14.1
Al ₂ O ₃ <250 µm	8.3	4.5	7.7	13.6
CaO < 2mm	2.6	0.4	2.2	6.0
CaO <250 µm	2.5	0.2	1.9	5.9
Fe ₂ O ₃ < 2mm	5.1	1.4	4.9	13.0
Fe ₂ O ₃ <250 µm	4.3	1.4	4.1	10.3
MgO < 2mm	0.9	0.2	0.8	1.8
MgO <250 µm	1.0	0.3	1.0	1.8
SiO ₂ < 2mm	60.7	32.0	62.0	78.0
SiO ₂ <250 µm	51.5	28.4	52.2	69.0

Table A-3 Summary statistics for <2mm and <250 µm fractions of topsoils from Swansea urban area (n = 25)

Variable	Mean	Minimum	Median	Maximum
B-As <250 µm	67	13	42	313
As < 2mm	313	55	170	2161
As <250 µm	320	53	171	2343
Al < 2mm	46129	26979	46023	73002
Al <250 µm	45198	25392	45494	67183
Ca < 2mm	19691	357	6718	102488
Ca <250 µm	15529	357	5360	73328
Fe < 2mm	49504	24549	41684	118828
Fe <250 µm	46323	26158	37977	132466
Mg < 2mm	4053	1809	3619	11459
Mg <250 µm	4005	1809	3619	12665
Si < 2mm	199035	106943	197074	282068
Si <250 µm	221134	128425	213886	316159

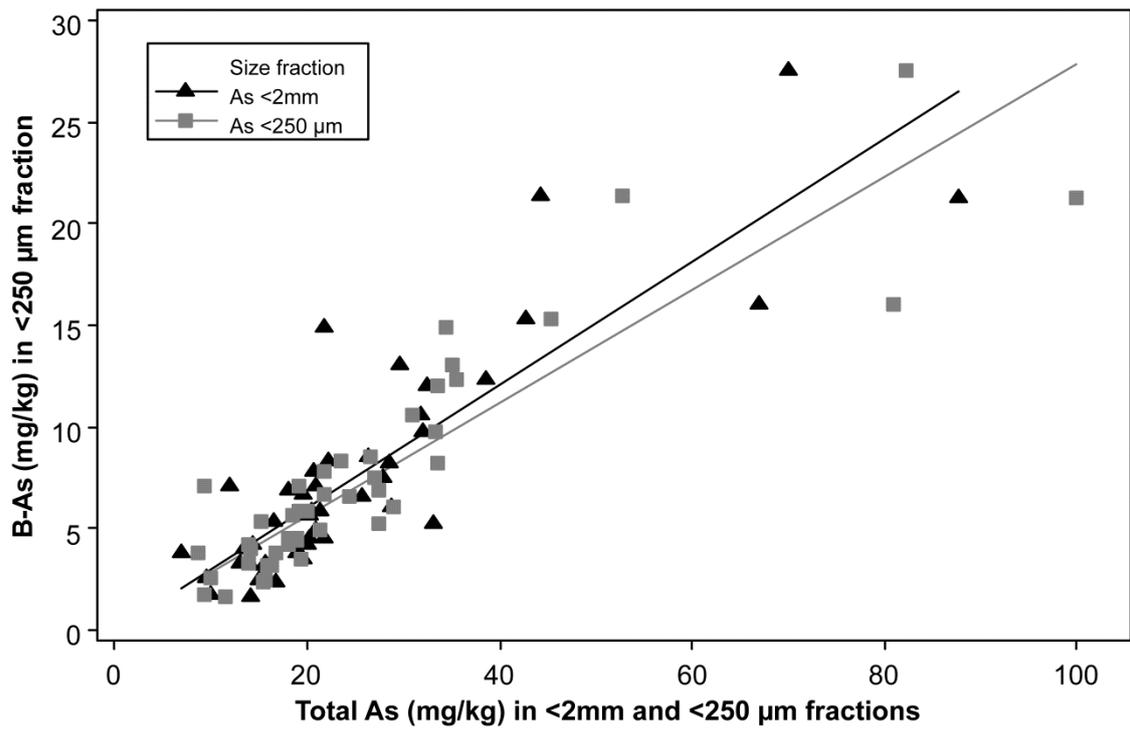


Figure A-4 Scatterplot for relationship between-As in <250 micron fraction and total As in <2mm and <250 micron fractions of London topsoils

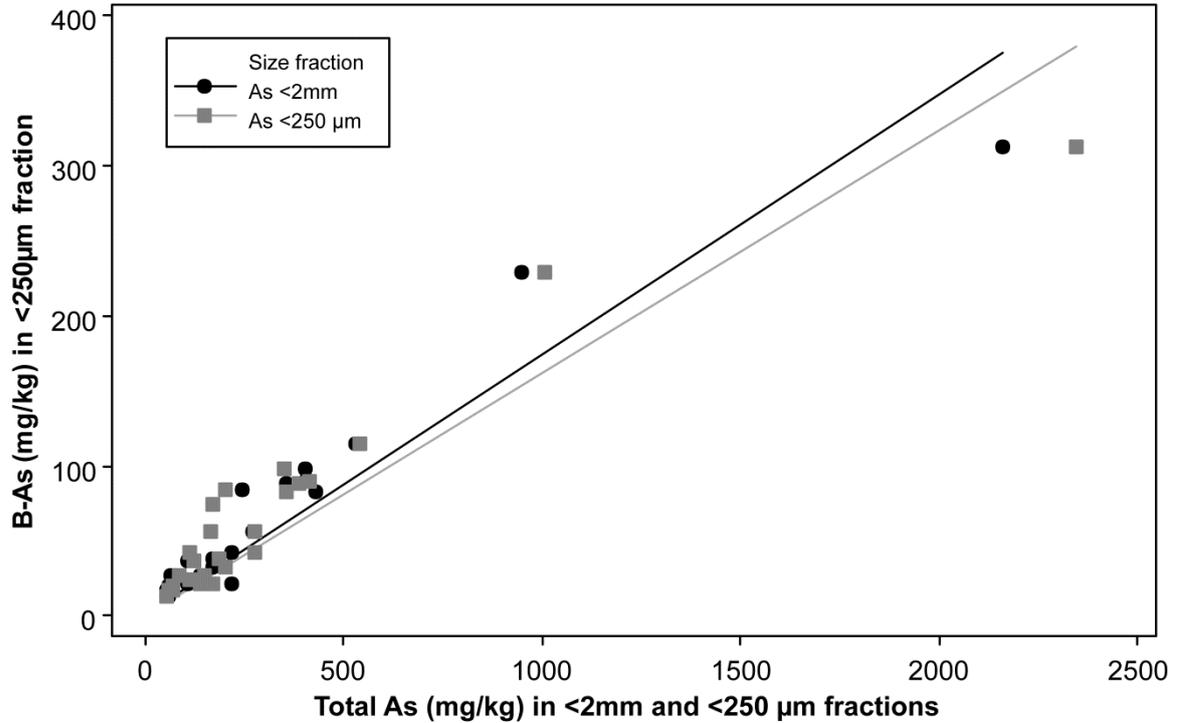


Figure A-5 Scatterplot for relationship between-As in <250 micron fraction and total As in <2mm and <250 micron fractions of Swansea topsoils