SUPPLEMENTAL INFORMATION

Non-targeted metabolomics and associations with per- and polyfluoroalkyl substances (PFAS) exposure in humans: A scoping review

Pengfei Guo[a,b], Tristan Furnary[a], Vasilis Vasiliou[a], Qi Yan[c], Kate Nyhan[a,d], Dean P. Jones[e,f], Caroline H. Johnson[a], Zeyan Liew[a,b,]*

[a] Department of Environmental Health Sciences, Yale School of Public Health, Yale University, New Haven, USA

[b] Yale Center for Perinatal, Pediatric, and Environmental Epidemiology, Yale School of Public Health, New Haven, USA

[c] Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles (UCLA), Los Angeles, USA

[d] Harvey Cushing / John Hay Whitney Medical Library, Yale University, New Haven, USA

[e] Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Department of Medicine, Emory University School of Medicine, Atlanta, USA

[f] Department of Biochemistry, Emory University School of Medicine, Atlanta, USA

*Corresponding Author
Zeyan Liew
Email: zeyan.liew@yale.edu
Address: 1 Church Street, New Haven, CT, 06510, USA

Contents

Appendix A. Search terms and database used.

Supplementary Table 1. Definitions of per- and polyfluoroalkyl substances (PFAS) mixtures from the studies reviewed.

Supplementary Table 2. Methodological and reporting quality of the studies reviewed.

Supplementary Table 3. A summary of statistical methods of the studies reviewed.

Supplementary Table 4. Analytical methods for PFAS exposure measurement.

Supplementary Figure 1. Scoring procedure for the reporting of exposure and metabolomics data.

Supplementary Figure 2. Venn diagrams of the commonly detected metabolic pathways associated with PFAS exposure by (A) sample type for metabolomics analysis, (B) study population, and (C) pathway analysis algorithm.

Abbreviations

Appendix A. Search terms and database used.

Dimensions:

(PFAS OR PFASs OR perfluoro* OR polyfluoro* OR perfluorinated* OR polyfluorinated* OR PFOS OR PFOA OR PFNA OR PFHxS OR PFSA OR PFCA OR PFOSA OR PFDA OR PFUnDA OR PFDeA OR PFDoA OR PFHpA OR PFUdA OR EtFOSAA OR MeFOSAA) AND (metabolome OR metabolomics OR metabolic OR metabonomics) AND (non-targeted OR non-target OR untargeted OR untarget OR mwas OR metabolome-wide)

Web of Science Core Collection as licensed at Yale:

TS=((PFAS OR PFASs OR perfluoro* OR polyfluoro* OR perfluorinated* OR polyfluorinated* OR PFOS OR PFOA OR PFNA OR PFHxS OR PFSA OR PFCA OR PFOSA OR PFDA OR PFUnDA OR PFDeA OR PFDo A OR PFHpA OR PFUdA OR EtFOSAA OR MeFOSAA) AND (metabolome OR metabolomics OR metabolic OR metabonomics) AND (non-targeted OR non-target OR untargeted OR untarget OR mwas OR metabolome-wide))

The databases linked through the Web of Science (WOS) as licensed at Yale were:

Web of Science Core Collection: Citation Indexes

- Science Citation Index Expanded (SCI-EXPANDED) --1900-present
- Social Sciences Citation Index (SSCI) --1900-present
- Arts & Humanities Citation Index (A&HCI) --1975-present
- Conference Proceedings Citation Index- Science (CPCI-S) --1991-present
- Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH) --1991-present
- Book Citation Index– Science (BKCI-S) --2005-present
- Book Citation Index– Social Sciences & Humanities (BKCI-SSH) --2005-present
- Emerging Sources Citation Index (ESCI) --2015-present

Web of Science Core Collection: Chemical Indexes

- Current Chemical Reactions (CCR-EXPANDED) --1985-present
  (Includes Institut National de la Propriete Industrielle structure data back to 1840)
- Index Chemicus (IC) --1993-present

Embase via Ovid

#	**Query**
1	(non-targeted or non-target or untargeted or untarget or mwas or metabolome-wide).mp.
2	(metabolome or metabolomics or metabolic or metabonomics).mp.
3	exp metabolome/
4	metabolomics/
5	(PFAS or PFASs or perfluoro* or polyfluoro* or perfluorinated* or polyfluorinated* or PFOS or PFOA or PFNA or PFHxS or PFSA or PFCA or PFOSA or PFDA or PFUnDA or PFDeA or PFDoA or PFHpA or PFUdA or EtFOSAA or MeFOSAA).mp.
6	exp perfluoroalkanoic acid/ or perfluorooctanesulfonic acid/ or perfluoro compound/
7	1 and (2 or 3 or 4) and (5 or 6)

Scopus

TITLE-ABS-KEY (((pfas OR pfass OR perfluoro* OR polyfluoro* OR perfluorinated* OR polyfluorinated* OR pfos OR pfoa OR pfna OR pfhxs OR pfsa OR pfca OR pfosa OR pfda OR pfunda OR pfdea OR pfdoa OR pfhpa OR pfuda OR etfosaa OR mefosaa) AND (metabolome OR metabolomics OR metabolic OR metabonomics) AND (non-targeted OR non-target OR untargeted OR untarget OR mwas OR metabolome-wide))) AND (EXCLUDE (SRCTYPE , "b") OR EXCLUDE (SRCTYPE , "re") OR EXCLUDE (SRCTYPE , "AGRI"))

Supplementary Table 1. Definitions of per- and polyfluoroalkyl substances (PFAS) mixtures from the studies reviewed.

Publication	PFAS mixture definition
Schillemans 2021	Two principal components (PC1 and PC 2) of six PFAS were derived using the principal component analysis. High loadings were observed for the long-chain PFDA, PFNA and PFUnDA in PC1 and for the shorter-chain PFHxS, PFOS and PFOA in PC2.
Chen 2020	n/a
Jin 2020	A principal component analysis of 3 PFAS (PFOA, PFOS, PFHxS) was conducted and the first component ("PC1"), which explained 73.2% of the variance, was used as a composite variable representing PFAS burden.
Li 2020	A PFAS community (PFOA, PFOS, PFHxS, EtFOSA) detected among the chemical variables from the correlation network, which was defined by all pairs of variables with Spearman correlation coefficient above 0.5.
Kingsley 2019	n/a
Hu 2019	n/a
Alderete 2019	A principal component analysis of 3 PFAS (PFOA, PFOS, PFHxS) was conducted and the first component ("PC1"), which explained 96.7% of the variance, was used as a composite variable representing PFAS burden.
Maitre 2018	n/a
Salihovic 2019	Total PFASs was constructed using the multivariable multiple regression (MMLR). All of the six PFASs (PFHxS, PFOS PFHpA, PFOA, PFNA, PFUdA) were included as independent variables and all metabolites were included as dependent variables. A significance testing was used to determine if the total metabolic profile differs between the PFASs.
Wang 2017	A molarity sum of all the detected PFCs, expressed as the total PFCs, including PFOS, PFOA, PFDA, PFNA, PFUnA, PFHxS, PFHxA, PFBA, PFDoA, PFHpA, and PFBS.
Lu 2019	Two kinds of total PFASs were used as mixtures: Σ6PFASs (PFBA, PFOA, PFBS, PFHxS, PFOS, 6:2 Cl-PFESA) and Σ13PFASs. The concentration of Σ6PFASs accounted for 99.7 ± 0.6% of the Σ13PFASs in the workers and for 88.6 ± 18.2% in the general population.

Supplementary Table 2. Methodological and reporting quality of the studies reviewed.

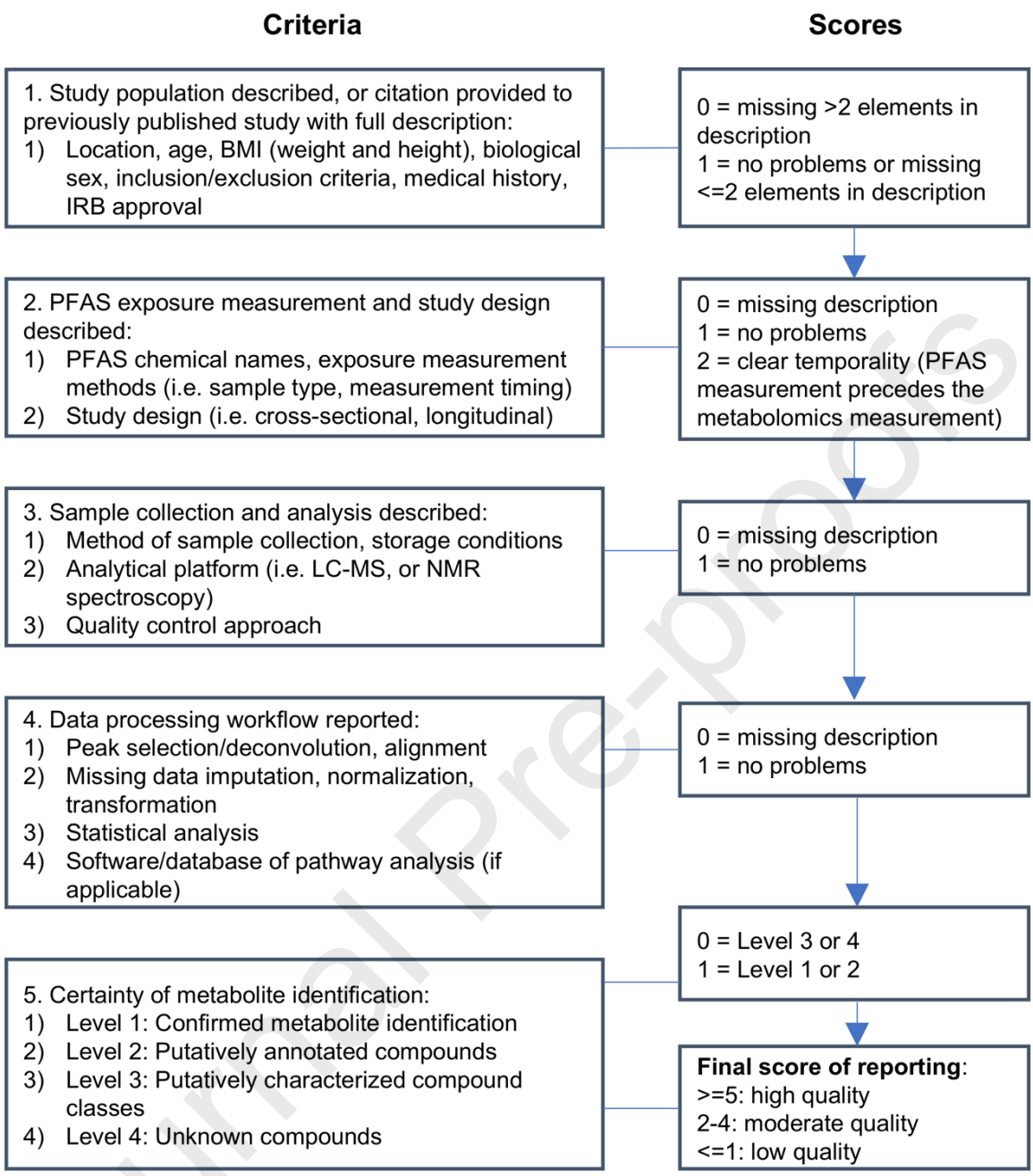
Publication	1. Study participants described	2. PFAS exposure assessment and study design	3. Sample collection and analysis	4. Data processing workflow	5. Certainty of metabolite identification	Total points
Schillemans 2021	1	1	1	1	1	5
Chen 2020	1	1	1	1	1	5
Jin 2020	1	1	1	0	0	3
Li 2020	1	1	1	1	0	4
Kingsley 2019	1	1	1	1	1	5
Hu 2019	1	1	1	0	1	4
Alderete 2019	1	2	1	0	0	4
Maitre 2018	1	2	1	1	1	6
Salihovic 2019	1	1	1	1	1	5
Wang 2017	1	1	1	1	0	4
Lu 2019	1	1	1	1	0	4

Supplementary Table 3. A summary of statistical methods in the studies reviewed.

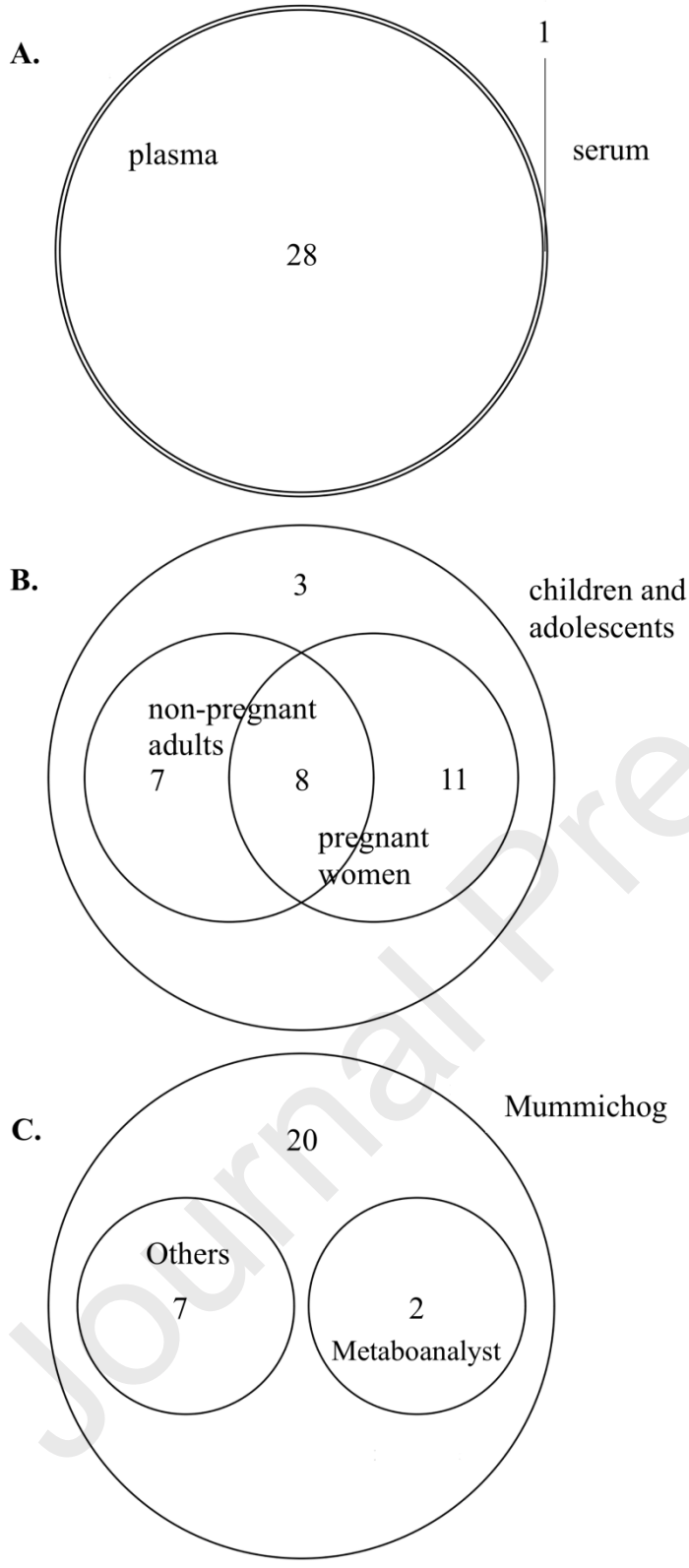
Publication	Statistical models employed to assess the associations between PFAS and metabolites	Covariates considered	Additional analyses
Schillemans 2021	Spearman's partial correlation (*i.e.*, covariate-adjusted Spearman's rank correlation)	sex, age, sample year, marital status, education, smoking status, physical activity and case-control status	n/a
Chen 2020	Linear regression	age, sex, parental education, race/ethnicity, cigarette and e-cigarette smoking status in the past week, physical activity levels and dietary covariates	The xMWAS integrated network analysis was used to investigate the associations among PFAS exposure levels, intensities of annotated metabolites, and cardiometabolic outcomes. The Sobel test for mediation analysis was conducted to evaluate the mediation effect of specific metabolites identified in the targeted metabolomic analysis.
Jin 2020	Linear regression	age, sex, ethnicity, and body mass index (BMI) Z-score	/
Li 2020	Hierarchical community network model	/	/
Kingsley 2019	Linear regression	child age, sex, and race	/
Hu 2019	Linear regression	total cholesterol, age, and p,p'-DDE levels	/
Alderete 2019	Linear regression	baseline age, sex, and social position (selected by directed acyclic graphs)	/
Maitre 2018	Spearman's partial correlation	gestational week, age and BMI	/
Salihovic 2019	Multivariable multiple regression (*i.e.*, modeling multiple dependent variables, with a single set of predictor variables)	all PFAS, sex, education level, smoking habits, exercise habits, energy intake and alcohol intake	Structural equation models were used to investigate a potential mediation of BMI on the relationship between PFAS and metabolites.
Wang 2017	Spearman's partial correlation	age, BMI, smoking and drinking status	/
Lu 2019	Spearman's partial correlation	age, BMI, gender, smoking, and drinking status	/

Supplementary Table 4. Analytical methods for PFAS exposure measurement.

Publication	Quality assurance and quality control procedure described	Lower limit of detection (LOD) (ng/mL)	Reported measured PFAS	Reported frequency of detection (%)	Statistical methods to address exposure values below the LOD
Schillemans 2021	Yes	0.027-0.049 (Koponen et al., 2013)	N/A	PFHxS, PFOS, PFOA, PFNA; PFDA (74), PFUdA (58)	Concentrations below the LOQ were set to the LOQ/2 = 0.075 ng/mL for 26% of the subjects in PFDA and 42% in PFUnDA
Chen 2020	Yes; CV<0.10	0.02-0.1	N/A	PFOA (100), PFOS (100), PFHxS (97.2)	N/A
Jin 2020	Yes (Go et al., 2015.)	0.02-0.1	N/A	PFOA (100), PFOS (100), PFHxS (94.6)	Set to a value equal to the detection limit divided by the square root of 2
Li 2020	Yes (Cohn et al., 2020)	0.028-0.11	11 PFAS	PFOS (100), PFOA (98), PFHxS (99), Others (1-50)	Set to missing
Kingsley 2019	Yes; CV~0.06 (Kato et al., 2014)	PFOS (0.2), PFNA (0.082), PFOA and PFHxS (0.1)	N/A	(100)	All serum PFAS concentrations were greater than the LODs
Hu 2019	N/A	N/A	11 PFAS	PFOS, PFOA, PFHxS, and EtFOSAA (>75), Others (<50)	N/A
Alderete 2019	Yes (Go et al., 2015.)	0.02-0.1	N/A	PFOA and PFOS (97.5), PFHxS (100)	N/A
Maitre 2018	N/A	N/A	4 PFAS	PFHxS (~55), PFNA, PFOA, PFOS	Set to the min value of the exposure divided by 2
Salihovic 2019	Yes (Salihovic et al., 2014)	0.01–0.17	12 PFAS	PFHxS, PFOS PFHpA, PFOA, PFNA, PFUdA, PFOSA, and PFDA (>75), Others (<75)	N/A
Wang 2017	N/A	0.06–0.14	11 PFAS	PFOA, PFOS, PFDA, and PFNA (100), PFUdA (95), PFHxS (58.6), Others (0)	N/A
Lu 2019	Yes (Gao et al., 2016)	0.008-0.19	21 PFAS	PFOS and PFOA (100), PFBA, PFBS, PFHxS, 6:2 Cl-PFESA, Others (0)	N/A



Supplementary Figure 1. Scoring procedure for the reporting of exposure and metabolomics data.



Supplementary Figure 2. Venn diagrams of the commonly detected metabolic pathways associated with PFAS exposure by (A) sample type for metabolomics analysis, (B) study population, and (C) pathway analysis algorithm. A total of 29 metabolic pathways associated with PFAS reported in at least three studies reviewed are included in the Venn diagrams. Each circle represents metabolic pathways associated with PFAS exposure in a specified group. The numbers listed in the circles represent the numbers of overlapping or non-overlapping pathways by study characteristics.