SUMMARY

Chronic Obstructive Pulmonary Disease (COPD) is a complex genetic disease characterized by progressive development of air-flow limitation, which encompasses chronic bronchitis and emphysema. Smoking and occupational air-borne dust are important etiological agents inducing an inflammatory response and are major risk factors in the development of COPD. Genetic susceptibility and environmental exposures triggering oxidative stress can further aggravate COPD pathogenesis. COPD is often under-diagnosed and mis-diagnosed because spirometry is not in routine use. Particulate matter (PM) and dust constituents at workplace (as at stone-crushing) sites can initiate inflammation and cause genetic damage and induce respiratory distress and eventually COPD. A case-control cross-sectional study as a first of its kind in an ethnic sub-group (SC migrants from Bihar) was hence planned to identify COPD cases in situ by spirometry at stone-crushing units and investigate them for genetic damage, oxidative stress and genotypes predisposing them to COPD. Unrelated healthy sub-group- matched adults from the general population comprised the control group. Approval for the study was obtained from the Institutional Ethics Committee, and cases (n=200) and controls (n=200) voluntary participating under informed consent formed the study group.

Cases were identified in situ from workers at stone-crushing units (n=22) located in the Pathankot and Gurdaspur districts of Punjab. Documentation on siting of units and compliance with regulations of the State and Central Pollution Control Boards was examined. Non-compliance was observed for: non-operational water sprinklers, absence of wind-breaking walls and proper sheds, non-metallic connecting roads, and non-provision of protective gear for workers. The top-soil layer in the 22 units contained PM (0.23-0.52μm) and 11 elements with silica and aluminium present in all. Using a face-to-face interview method on a specially-designed proforma, information on general demographics, life-style, work-related information, exposure history, family history and reproductive performance (pedigree) of the study group was recorded. In situ measurements were taken using standard protocols for some anthropometric (height, weight, waist circumference and hip circumference) and physiometric (systolic and diastolic blood pressure (SBP/DBP) variables. Pulmonary function test by spirometry
assessments using the Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) guidelines was performed by 258 workers and 200 (77.51%) were identified as having chronic obstructive pulmonary disease. A non-exposed healthy group matched for age, gender, socio-economic status, life-style patterns and ethnicity from the general population with normal pulmonary function comprised the control group (n=200).

Sample collections were made as per standard protocols for venous blood (~10ml), buccal mucosa and urine (~200 ml). Genomic DNA investigations comprised assessment of peripheral blood leukocytes (PBL) for DNA damage (after cell viability test) using the alkaline Single Cell Gel Electrophoresis/SCGE (comet) assay and of oxidative DNA damage by the enzymatically-modified comet assay. Buccal epithelial cells were processed for the Buccal Micronucleus Cytome (BMNCyt) assay for chromosomal damage (MN-micronuclei from aneugenic/clastogenic events), DNA damage (nuclear buds from DNA amplification), cytokinetic defects (binucleated cells), cell-proliferation (basal cells), and cell-death markers (karyorrhectic, condensed chromatin, karyolytic and pyknotic cells). The urothelial cells were analysed for chromosomal damage (MN) and cell-death events (karyorrhectic, karyolytic, condensed chromatin and pyknotic cells) by the urothelial MN test. Blood-sera samples were assessed for oxidative stress (total antioxidant capacity-TAC, total oxidant status-TOS and lipid peroxidation viz. Malondialdehyde (MDA) levels). Biochemical analysis also comprised estimation of alpha-1-antitrypsin (AAT) activity and of lipid levels. Whole blood genomic DNA was genotyped for seven SNPs comprising four variants of one candidate gene, SERPINA1 (rs 6647 (710T>C), rs 709932 (374G>A), rs 17580 (863 A>T), rs 28929474 (1096G>A)) and three metabolic gene variants of GSTT1, GSTM1 and GSTP1 (313A>G) by the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis.

Statistical analyses included comparison between cases and controls for demographic and life-style patterns, levels of genetic damage, oxidative stress, AAT, lipids and lipoproteins as well as of allelic and genotypic distributions using appropriate tests. Specificity and sensitivity of assays were determined by drawing the Receiver Operating Characteristic (ROC) curves. Principal Component Analysis (PCA) was performed for factor reduction and Multidimensionality Factor Reduction (MDR) analysis for gene-gene and gene-environment interactions. Association analysis and
identification of predictors for genetic damage and disease-risk were also carried out. Likelihood ratios were calculated for identifying ‘at-risk’ genotypes. The overall retrospective power of the study considering all study variables ranges between 82-90% at 5% level of type I (α) error.

COPD cases (41.69±0.47y; 28-64y) and controls (41.58± 0.58y;25-58y) were all male migrants from the state of Bihar and were matched for dietary habits and smoking duration but differed for smoking frequency (heavy smokers among cases) and BMI (more underweight cases). Other differences were observed for physiometric analysis (with more pre-hypertensive and normotensive cases), general obesity (more underweight cases, BMI basis) and for waist circumference (WC) and waist-hip- ratio (WHR, more obese controls).

Occupational duration at stone-crushing units was of 12.87±0.24y (5-21y) with daily work shifts of 11-13h (9.83±0.14h) and involvement in activities of drilling (35.50%), dressing (34.00%) and loading (30.50%) of stones. Controls were mainly employed at dairies, cycle-repairing shops at roadsides and tuck-shops (dhabas).

Pulmonary function defining COPD among cases exhibited significant (p=0.000) decrease in forced expiratory volume (1.55X) in one second (FEV₁), in expiratory ratio (FEV₁/FVC) by 1.55x, in per cent FEV₁ predicted by 1.50x and of body surface area (BSA) by 1.05x in cases compared to values in controls. GOLD classification defined moderate (34.00%), severe (41.50%) and very severe (24.50%) COPD cases. Respiratory distress symptoms were reported by all cases and comprised shortness of breath (96.00%), persistent cough (48.00%), blocked nose (94.50%) and dry/sore throat (93.50%). Other ill-health symptoms included eye-irritations (94.00%), eye-watering (89.50%), fever (83.50%), fatigue (97.50%), joint/muscle pain (92.00%), abdominal pain (48.50%), loss-of-appetite (43.00%) and worms in faeces (37.50%). Pedigree analysis revealed significantly more number of abortions, miscarriages and still births in cases but no birth defects were reported. Also there was no COPD family history or inter-caste marriages in the three-generational pedigrees of cases and controls.

DNA damage and the oxidative DNA damage were both significantly increased in cases viz. per cent DNA in tail (1.08x,p=0.000), damage index (DI-1.07x,p=0.000),
damage frequency (DF, p=0.034), oxidized pyrimidines (1.43x,p=0.000), oxidized purines (1.42x,p=0.000) and total oxidative damage (1.43x,p=0.000). There was more oxidative damage to pyrimidines than to purines (1.38x,p=0.000).

The BMNCyt assay revealed significant increase of micronucleated cells (1.16x, p=0.000), nuclear buds (1.40x,p=0.000) and the cell-death parameters comprising karyorrhectic (1.14x,p=0.000), condensed chromatin (1.96x,p=0.000) and karyolytic (1.07x,p=0.006) cells. Cell-kinetic defects were increased by 2.75x (binucleated cells) and cell-proliferation (basal cells) was decreased (1.07x). Repair index which implies genotoxicity was also significantly (1.06x,p=0.030) decreased in cases compared to the controls. In the urothelial MN assay, chromosomal damage (micronucleated cells,1.94x) and the atypical cell-types viz. karyolytic (1.47x) and condensed chromatin cells (2.75x) cells were also significantly (p=0.000) increased in cases. Oxidative stress in cases was elevated compared to levels in controls as indicated by significantly decreased TAC (1.23x,p=0.030) and increased TOS (1.31x,p=0.000), oxidative stress index (OSI 3.11x,p=0.000) and MDA (1.38x,p=0.000). Significant decrease in activity of AAT (1.21x,p=0.000) and levels of total cholesterol (TC 1.09x,p=0.000), high density lipoprotein (HDL, 1.11x,p=0.014) and low density lipoproteins (LDL 1.15x,p=0.001) were also observed. Atherogenic indices of LDL/HDL (1.28x,p=0.013), TC/HDL (1.27x,p=0.002) and TG/HDL (1.52x,p=0.01) were significantly raised implying a significant cardiovascular disease-risk among cases.

Disease severity viz. severe, very severe and moderate COPD cases were compared with each other for demographic, clinical and exposure-related variables and also the two former with controls. Comparison with controls revealed matching for diet, number of bidis smoked, Brinkmann index, workplace activities, per day schedule of work, blood pressure measurements, Body Mass Index, waist circumference, waist-hip-ratio as well as the waist-height-ratio. However, the very severe cases were significantly older (>45y), had been smoking bidis for significantly more duration (>20y) and had worked for more duration (>15y) at stone-crushing units than the moderate and severe cases. The lung-function parameters declined significantly (p=0.000) with severity. Respiratory distress was significantly more in very severe cases followed by that in severe and then in moderate cases.
On comparison of the assessed parameters for disease-severity, the micronucleated cell frequency (p=0.000) in buccal epithelial cells, repair index (p=0.000) and the lipid peroxidation levels (p=0.002) were significantly increased in severe and very severe cases (but not between them) in comparison to the moderate COPD cases. Other genetic damage parameters and the oxidative stress and biochemical markers did not differ significantly between the three groups. Increased severity was related with workplace exposure, smoking history, anthropometric characteristics but did not show any trend.

Stratification of the data for differential genetic damage, oxidative stress, AAT levels and lipid profile was carried out as a function of age (≤45y versus >45y), study sites (Gurdaspur vs. Pathankot), workplace exposure (≤10h versus > 10h duration in occupation, ≤10h versus > 10h daily shift), workplace activities (drillers vs dressers vs loaders), smoking history (no. of bidis smoked per day, duration of smoking history, Brinkmann index), physiometric characteristics (optimal vs pre-hypertension vs Hypertension Stage I vs Hypertension Stage II general, obesity (BMI)), central/abdominal obesity (WC, WHR, WHtR).

Cases with longer working duration had significantly increased micronucleated (1.15x,p=0.000) buccal cells and condensed chromatin (1.61x,p=0.000) urothelial cells, and damage index (1.04x,p=0.034) in peripheral blood leukocytes. Drillers compared to dressers had increased condensed chromatin (1.38x,p=0.040) urothelial cells. Dressers compared to loaders had significantly higher levels of TC (1.07x,p=0.000), HDL (1.06x,p=0.023) and LDL (1.06x,p=0.001) but decreased atherogenic indices (TC/HDL, 1.06x, p=0.046;TG/HDL, 1.23x,p=0.006).Increased smoking duration revealed significantly elevated DNA damage (micronucleated cells, 1.13x,p=0.000; nuclear buds (1.07x,p=0.000) and cytokinetic defects (binucleated cells, p=0.038) in buccal cells. Also with increased Brinkmann index (BI) , there was significant increase in oxidative DNA damage (1.53x,p=0.002 in peripheral blood leukocytes), of micronucleated (1.16x,p=0.021) buccal cells and condensed chromatin (1.68x,p=0.042) urothelial cells. Underweight cases as per BMI had significantly increased frequency of binucleated cells (1.26x,=0.001) in buccal samples and karyolytic (1.88x,p=0.014) urothelial cells in comparison to those with normal BMI values. Also obese cases on WHtR basis had elevated damage index (1.05x,p=0.002) compared to levels in non-obese cases. Association (ANOVA, correlation analyses) of diet, BMI, WC, WHtR, smoking history
and FEV₁/FVC was revealed for genetic damage (in peripheral blood leukocytes, buccal mucosa, urothelial cells), oxidative stress, AAT activity and lipid profile. Among soil constituents, silica, aluminum, silver, sodium, magnesium, potassium and iron emerged as significant predictors for genetic damage, and silica, aluminum and sodium of oxidative stress and AAT activity as well.

Molecular genotyping results revealed Hardy-Weinberg equilibrium for all SNPs, except of \textit{SERPIN A1} (rs17580, A863T; rs 28929474, G1096A) which were all wild type homozygotes in cases and controls. Genotype differences were revealed only for the distribution of \textit{GSTM1} with null genotype which was significantly (p=0.045) higher in cases. The minor allele frequency of \textit{GSTM1} was higher in cases (0.72 vs 0.64 in controls) and the additive inheritance model was the best-fit, conferring a significant risk (p=0.0358) for COPD.

Stratification for disease severity further revealed that heterozygous (TC) and homozygous (CC) variants of \textit{SERPINA1} (rs6647; 710T>C) and the homozygous (AA) variant of \textit{SERPINA1} (rs 709932; 374 G>A) significantly (p=0.000) conferred risk for severe COPD. Among combinational genotypes (2-,3-,7-SNPs), the two-SNPs’ combination of \textit{GSTT1} present and TC genotype of \textit{SERPINA1 M1} (OR 0.529, 95\%CI 0.328-0.85, p=0.009) and three- SNP combination of \textit{GSTT1} present, \textit{GSTM1} null and AA (313A>G) genotype of \textit{GSTP1} (OR 1.81, 95\%CI 1.04-3.14, p=0.033), and the 7-SNPs’ combination of \textit{GSTT1} present, \textit{GSTM1} present, \textit{GSTP1(AA)}, \textit{SERPINA1 S} (AA), \textit{Z} (GG), \textit{M2} (GA) and \textit{M1} (CC) (OR=2.01,95\%CI 1.06-2.14,p=0.023) emerged as disease-risk combinational genotypes.

Genotypic distribution in context of genetic damage and other parameters in cases and controls revealed modulation by the \textit{GSTT1} and \textit{SERPINA1 M1} and \textit{M2} variants. The \textit{GSTT1} present genotype had significantly increased micronucleated cells (p=0.053) and VLDL levels (p=0.010) and decreased LDL levels (p=0.025). For the \textit{SERPINA1 M1} (710T>C) the homozygous (CC) variant had significantly (p=0.001) increased frequency of condensed chromatin cells as also the (CT) genotype (p=0.003) in comparison to the frequency in homozygous (TT) variant genotypes. Also the CC versus CT had significantly increased TC (p=0.023) and LDL (p=0.045) levels though LDL/HDL (p=0.022; 0.006) and TC/HDL (p=0.032; p=0.033) levels were significantly
increased in homozygous (TT) genotypes of SERPINA1 M1 in comparison to the homozygous (CC) variant and heterozygous (CT) genotypes. For SERPINA1 M2 (374G>A) cases with AA vs GG genotypes, the former had significantly increased percent DNA in tail (p=0.021), damage index (p=0.030), oxidized pyrimidines (p=0.001), oxidized purines (p=0.001) and total oxidative damage (p=0.000). The karyolytic cell frequency in buccal cells was highly significantly increased, both in GG (p=0.050) and AA (p=0.038) genotypes in comparison to those with heterozygous (GA) SERPINA1 M2 genotype.

Data reduction analysis (PCA) was performed for disease-risk and revealed 80.80 % variance; PCA for genetic damage exhibited 66.70% variance while maximum variance of 93.61% was revealed for genotypic data. For disease-severity PCA, with 80.80 % variance 27 factors were loaded. The factor one loading comprised age, number of bidis smoked, years since smoking, Brinkmann index and COPD severity with 9.61% of the variance, factor two loaded with silica and karyorrhectic cells in buccal cells accounting for 15.71% variance, factor three loaded with FEV1, FVC, % FEV1 predicted, %FVC predicted and COPD severity accounting for 21.46% variance. Besides other factors with various predictors, the factors six loaded (with GSTM1 null) genotype and factor nine GSTT1 null (variance of 45.98%). For genetic damage, PCA revealed eight factor loadings with 16.43% variance and factor one comprised net oxidative DNA damage, factor two with binucleated cells and repair index in buccal cells, factor three for percent DNA in tail and DI, factor four with binucleated cells and factors 5-7 with atypical urothelial cells.

Among genotypic data on PCA, 15 factors were loaded with 19.28% variance and the GSTT1 loaded on factor one, GSTP1 heterozygous genotype on three, SERPINA1 M2 heterozygous genotype on four, SERPINA1 M1 heterozygous genotypes on five and SERPINA1 M1 homozygous variant genotype on factor six.

Gene-gene interactions (MDR analysis) for disease-risk revealed that the trivariate model comprising GSTT1, SERPINA1 M1 (710 T>C), SERPINA1M2 (374G>A) exhibited 2.17 times higher (OR 2.17; 95% CI 1.45-3.24; p=0.0001) disease-risk with CV consistency of 9/10 and with CV balance accuracy of 0.535. This trivariate model further showed different genotype combinations of GSTT1 (null/present) and SERPINA1 M1 and M2 as ‘high-disease-risk’ genotype-combinations.
Gene-environment interactions for disease-risk revealed that three models viz. univariate comprising work-activities, bivariate comprising \textit{GSTT1} with work-activities, and the trivariate model comprising \textit{GSTT1, GSTM1} and the work-activities. The latter model conferred a significant (p=0.001) risk for disease with CV consistency of 10/10 and with CV balance accuracy of 1.0. The models revealed varied combinations of drillers/dressers/loaders/GSTT1 (null/present)/GSTM1 (null/present) as ‘at-risk’ individuals for COPD.

Receiver Operating Characteristics (ROC) curve analysis underlying the sensitivity (false negatives) and specificity (false positive) of the three genetic damage assessment assays revealed that the buccal micronucleus assay had the highest (96.50%) specificity and sensitivity and was the most valid biomarker followed by the SCGE Assay (85.50% specificity, 83.00% sensitivity) and the Urothelial MN (82.00% specificity, and 80.00% sensitivity) assay. As these three assays are well- validated for population bio-monitoring, the multitude biomarkers assessed in each assay were also subjected to ROC analysis. Among the SCGE assay, damage frequency (85.50%) and per cent DNA in tail (82.00%), for the BMNCyt assay, binucleated (98.50%), micronucleated (92%) and condensed chromatin (96.00%) cells; and among urothelial cells, the karyolytic (91.50%) and condensed chromatin (93.00%) cells showed significant specificity.

Plausible evidence for disease- etiology lies in the presence of particulate matter (PM 0.23-0.53\textmu m) and carbon, oxygen, silica, aluminium, calcium, magnesium, sodium, copper and iron present in the top-soil layer of all units (detectable by Energy Dispersive X-ray Spectroscopy with Scanning Electron Microscope -EDX-SEM) to which those working at the units are exposed. Both, the presence of PM, silica and aluminium in dust-emissions are known COPD-etiological factors.

Hence, the study out.puts highlight undiagnosed cases of COPD as an occupational disease from exposure to PM dust and various elements at stone-crushing units, with \textit{GSTT1, M1} and \textit{SERPINA1 M1,M2} genotypes as at-risk genotypes, with a susceptibility to increased genomic damage scored in three different biological samples along with the presence of oxidative stress in blood serum.