Introduction

Recovery patterns from work-related musculoskeletal disorders (WRMSDs) vary considerably. These can be manifested by unresolved pain, decrements in physical functioning, persistent symptoms and lost work time [1–3]. Factors affecting recovery involve complex interrelationships among personal and workplace (physical and organizational) influences [4–9]. The links between individual-level socio-economic status (SES) and work factors are particularly relevant to post-injury recovery and may be interrelated, given that a component of SES-related gradients in health can be explained by workplace physical and organizational variables [10–13].

The present study was designed to assess incident WRMSDs and recovery in hospital workers, including both cross-sectional and longitudinal components. We have previously reported details of our study design and cross-sectional findings of incident WRMSD linked to SES and job factors [14–17]. Another study of hospital...
workers subsequently also reported that injury was associated with SES and physical and organizational work factors [18].

We wished to analyse baseline injury status as well as SES and physical and organizational work factors as predictors of health and work status at 2 year follow-up. In particular, we aimed to ascertain whether health deficits associated with WRMMSD at baseline (defined by body region-specific functional limitations and general health status) resolved over time. We also wished to determine whether work productivity and work status were negatively affected at follow-up by baseline injury, SES and work factors.

**Methods**

We linked baseline and 2 year follow-up data obtained in the Gradients of Occupational Health in Hospital Workers (GROW) study. The longitudinal component of this investigation followed injured workers (cases) with incident WRMSDs of the trunk, neck and upper and lower extremities from two separate hospital sites comparing them to non-injured referents from the same sites. The protocol was approved by the University of California San Francisco committee on research involving human subjects.

We collected baseline data in 2002–04 through structured telephone interviews employing computer-assisted telephone interviewing software. In addition, we conducted onsite ergonomics observations of work practices in a subset of subjects (75%). Details of the study design, recruitment and validity and reliability of the baseline interview and ergonomics instruments used in the subset analysis as well as the baseline study findings have been previously published [14–17]. In brief, we recruited participants from a base of ~6000 hospital workers at two sites, representing all occupational groups (with the exception of physicians who were excluded from the study). Cases were defined by an incident WRMMSD determined to be work related by physicians or nurse practitioners employed at each site’s employee health clinic. Referents were equally matched on the basis of (i) job group, (ii) shift work type (e.g. working a routine daytime schedule compared to various shift arrangements) or (iii) at random sequentially in time (incidence density matching), yielding an overall 3:1 ratio of referents to cases. Approximately 2 years after baseline participation, we attempted to recontact subjects for follow-up structured telephone interviews.

Data for age, sex, race–ethnicity, education, income, smoking status and medical co-morbidities were obtained at baseline. Occupational categories were grouped as administrators and managers, nursing, other clinical, clerical, technical or support staff. We assessed work organization factors using two measures: (i) job strain, derived from the Job Content Questionnaire (JCQ) [19] and (ii) the effort–reward ratio, derived from the Effort-Reward Imbalance (ERI) Questionnaire [20]. Three ergonomics measures, based on direct worksite observations, were assessed at baseline [16]. The first ergonomic measure assessed upper body and neck strain [Upper Body Assessment—University of California (UBA-UC)]; the second, back and lower extremity strain (LBA-UC) and the third, the observed proportion of time spent using a computer. These measures were summarized for all individuals within each of 13 job categories within the study population and then applied to all individuals within that category [16].

General health status was measured at both baseline and follow-up using the physical component summary of the Short-Form 12 (PCS) [21], which theoretically ranges from 0 to 100 (higher scores reflecting better functional status). Body region-specific health status instruments assessing disability and pain were administered to cases and matched referents for each of four injury types [back, upper extremity (UE), lower extremity (LE) and neck]. Low back symptoms and pain were assessed using the Roland–Morris Scale [22], which ranges from 0 to 24 (higher scores denote worse functional status). UE symptoms were measured using the 11-item Quick Disability of the Arm, Shoulder and Hand (DASH) instrument, which ranges from 0 to 100 (higher scores represent more severe and disabling symptoms) [23]. The severity of LE symptoms was assessed using a shortened version of the Western Ontario and McMaster Universities (WOMAC) osteoarthritis index [24], which ranges from 0 to 100 (higher scores indicate better functional status).

Two work status measures (work effectiveness and lost workdays) were evaluated. Work effectiveness was assessed using a self-reported work effectiveness score, ranging from 0 to 100% (0% corresponding to inability to work at all and 100% indicating greatest effectiveness). Lost workdays for any cause in the 4 weeks preceding the interview were also elicited. We also ascertained whether subjects were no longer working at their original job site or at any job.

To assess whether injured cases had regained functioning by 2 years relative to all referents, we compared scores for cases versus referents at baseline and at follow-up for all measures assessed at both time points (PCS, work effectiveness, lost workdays and injury-specific disability). Due to the non-normal distribution of these results, we used the Wilcoxon rank-sum test to test differences by injury status (Figures 1 and 2). Because statistical differences between continuous scales may be of marginal clinical relevance, we created a binary measure for each of these outcomes dichotomizing between poor and relatively better health or work status. For most outcomes, the threshold for poor functioning was determined based on the quartile distribution for each score among
referents. For measures in which higher scores reflect better status (PCS and WOMAC), poor status was based on the lowest quartile; for measures in which higher scores reflected worse functioning (Roland–Morris and DASH), the highest quartile defined poor status. The threshold for self-rated work effectiveness was set to 90%, consistent with our previous dichotomization of this measure [15]. For lost workdays, we defined poor status as two or more lost workdays in the past 4 weeks, a cut point approximating the 37th percentile.

Univariate logistic regression tested whether injury status was associated with poor functional status at baseline or at follow-up as well as the association between injury and changing to a job at another location (whether or not this involved job duty changes) or complete work cessation. After a screening step based on univariate analyses employing a statistical significance cut-off of $P < 0.05$ for other baseline factors of interest (demographics, occupational category, clinical characteristics, job strain defined by the JCQ, ERI and ergonomics measures), we used multivariable logistic regression to ascertain (i) the degree of association between injury status and poor functional status, including adjustment for baseline cofactors of interest, and (ii) whether other baseline factors of interest were predictive of poor functional status at follow-up taking injury status into account. Age, sex and race–ethnicity were retained in the models regardless of statistical significance at the univariate screening step due to their known influence on health and functioning.

Because of collinearity between the education and income variables, we combined them for the regression analysis by adding one point each for higher levels of each and grouping that sum into quintiles, a measure of SES we had employed in the baseline analysis [15]. Similarly, we combined data from the collinear JCQ and ERI measures by reducing each into a binary above-median or below-median dichotomous variable, creating four mutually exclusive indicator variables: above the median for both

---

**Figure 1.** Comparison of case and referent general health and work status distributions at baseline and follow-up. *Wilcoxon $P < 0.05$; **Wilcoxon $P < 0.01$; ***Wilcoxon $P < 0.001$. DASH, Disability of the Arm, Shoulder and Hand; WOMAC, Western Ontario and McMaster Universities osteoarthritis index.
(worst quadrant), above the median for one but not the other (middle quadrants) and below the median for both (the default referent category).

We also investigated whether lowest quintile of SES or high combined job strain/effort–reward imbalance, which we hypothesized a priori could be effect modifiers for injury status, should be included in our final models. This was accomplished by rerunning the key multivariate models including each interaction term separately. Interactions that were statistically significant at $P < 0.10$ and did not include sparse cells (sample size $n < 10$) were considered for inclusion.

We imputed missing data for individuals without any follow-up information ($n = 70$) and for subjects missing one or more key dependent or independent variables on either the baseline or follow-up assessments ($n = 95$). This imputation was conducted using multiple imputation procedures (SAS version 9.1.3 PROC MI and PROC MIANALYZE). Specific variables imputed were family income at baseline, education at baseline and all outcome measures.

### Results

We reinterviewed 582 (88%) of the 664 baseline subjects (median time elapsed 1.9 years (659 days) 25–75th percentile 575–724 days). Baseline characteristics by follow-up status are presented in Supplementary Table 1 (available as Supplementary data at Occupational Medicine online). Those not reinterviewed manifest a different racial–ethnic mix and lower levels of educational attainment ($P < 0.001$ in both instances). Follow-up status also differed overall by job group ($P < 0.01$).

Figure 1 displays a comparison of general and injury-specific health status measures at baseline and follow-up for injured cases and referents. The PCS showed significantly lower values for cases versus referents at baseline [median 39 versus 52 ($P < 0.001$)]; this difference narrowed, but was still statistically significant, at follow-up [median 49 versus 52 ($P < 0.01$)]. PCS values among the cases, even though shifted, continued to display a modest bimodal distribution. Work effectiveness was significantly lower for cases versus referents at baseline.
[median 90 versus 98% ($P < 0.001$)], but by follow-up, this difference was negligible (median 99% versus 98% and no longer statistically significant) and the overall distributions were quite similar. The distribution of lost workdays in the 4 weeks preceding interview differed statistically between groups ($P < 0.001$) at baseline; once again, by follow-up, there was no longer a substantive or statistically significant difference in the distribution of lost workdays.

Distribution of body region-specific health status outcomes for the strata of cases by injury type and their matched referents are shown in Figure 2. The Roland–Morris scores for back injury cases were statistically significantly worse for cases at baseline [median 10 versus 0 ($P < 0.001$)]; at follow-up, this gap had narrowed, although it was still statistically significant; further, the distribution shows that few cases at follow-up reported very low Roland–Morris scores indicative of no limitation whatsoever. Similarly, DASH scores for UE injury cases, which reflected significantly worse functioning compared to referents at baseline [median 34 versus 6 ($P < 0.001$)], also attenuated, but a gap remained [median 14 versus 8 ($P < 0.01$)]. Baseline differences for the LE, using the WOMAC, were significantly decreased for cases compared to referents [median 45 versus 100 ($P < 0.01$)], but this difference was no longer statistically significant at follow-up, and the distributions of scores for cases and referents were quite similar.

Table 2 shows the injury-associated risk of poor PCS and body region-specific limitations (baseline and follow-up) and work-related outcomes (follow-up only). Baseline WRMSDs were associated with increased risk of poor PCS at baseline [odds ratio (OR) = 5.2; 95% confidence interval (CI) 3.5–7.5] and at follow-up (OR = 1.5; 95% CI 1.0–2.3; $P < 0.05$). All the body region-specific health status measures showed significantly higher likelihoods for poor status at baseline; these associations were all attenuated at follow-up and none was statistically significant at the $P < 0.05$ level. Injury cases manifested 40–50% increased odds of no longer being

### Table 2. Injury case status as a predictor of poor functioning at baseline and follow-up within the GROW Study Cohort

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Baseline</th>
<th></th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>Cases with outcome</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td></td>
<td>Cases/Ref</td>
<td>$n$ (%)</td>
<td></td>
</tr>
<tr>
<td>SF-12 physical component score, lowest quartile</td>
<td>166/498</td>
<td>103 (62)</td>
<td>5.2 (3.5–7.5)</td>
</tr>
<tr>
<td>Work effectiveness, lowest quartile</td>
<td>166/498</td>
<td>91 (55)</td>
<td>2.2 (1.6–3.2)</td>
</tr>
<tr>
<td>Lost workdays (2+ in last four weeks)</td>
<td>166/498</td>
<td>79 (48)</td>
<td>1.5 (1.1–2.2)</td>
</tr>
<tr>
<td>Roland–Morris (back injury), highest quartile</td>
<td>50/150</td>
<td>31 (62)</td>
<td>5.8 (2.9–11.5)</td>
</tr>
<tr>
<td>DASH (UE), highest quartile</td>
<td>74/222</td>
<td>51 (69)</td>
<td>6.8 (3.8–12.2)</td>
</tr>
<tr>
<td>WOMAC physical functioning (lower extremity), lowest quartile</td>
<td>24/72</td>
<td>11 (46)</td>
<td>2.9 (1.1–7.8)</td>
</tr>
<tr>
<td>No longer working at original job site</td>
<td>_b</td>
<td>_</td>
<td>_</td>
</tr>
<tr>
<td>No longer working at any job</td>
<td>_b</td>
<td>_</td>
<td>_</td>
</tr>
<tr>
<td>Work status change (left original job site or no longer working)</td>
<td>_b</td>
<td>_</td>
<td>_</td>
</tr>
</tbody>
</table>

Poor functioning defined as lowest functioning quartile over both assessments for the referents subsample for SF-12 PCS and work effectiveness, Roland–Morris, DASH and WOMAC. For lost workdays, poor functioning defined as ≥2 days lost in the past 4 weeks. Ref, Referents.

_aN_ for outcomes applicable only to individuals working at follow-up is based on rounded number of individuals with an actual or imputed value of working at follow-up. $N = 13$ cases and $n = 29$ controls were not working at follow-up, which due to rounding error does not add up to the $N = 43$ who were not working when cases and referents are pooled.

Only applicable at follow-up.
employed at their original worksite or not working at all, although this association was not statistically significant (for worksite change or work cessation, OR = 1.5; 95% CI 0.9–2.4).

Supplementary Table 3 (available as Supplementary data at Occupational Medicine online) presents an analysis of the other potential baseline predictors of poor PCS and adverse occupational outcomes at follow-up. Poor PCS and a change in work status (defined by site change or work cessation) were significantly associated with lowest quintile SES (PCS: OR = 2.2; 95% CI 1.2–4.1 and site change/not working: OR = 2.4; 95% CI 1.1–5.1, respectively). Subjects at baseline above the median for job strain/effort–reward imbalance also manifested another adverse outcome: reduced work effectiveness at follow-up (OR = 2.5; 95% CI 1.6–4.0). No other variable was associated with more than one of the outcomes studied. None of the three baseline ergonomics measures analysed was associated with more than one of the outcomes studied. The median values for job strain (JCQ)/ERI. These adjusted ORs are similar to the unadjusted ORs displayed in Table 2, although the 95% CIs are wider and no longer excluded 1.0 for any measure. For poor PCS, the injury-associated risk estimate was slightly attenuated: the OR was reduced from 1.5 to 1.4 (adjusted 95% CI 0.9–2.2). None of the covariates adjusters was consistently associated across all the outcome measures. The significant relationships noted in the univariate regressions remained, with the exception of ‘other’ race that was no longer a significant predictor of work status change. Lowest quintile of SES remained predictive of both poor PCS (OR = 2.0; 95% CI 1.0–4.0) and work status change (OR = 2.5; 95% CI 1.1–5.8). The relationship of SES categories to these outcomes was not monotonic, with increased risk of lower PCS also associated with the next to highest quintile; this relationship was of borderline significance (OR = 1.8; 95% CI 1.0–3.4). Combined above-median job strain/effort–reward imbalance at baseline remained associated with two adverse outcomes: poor PCS (OR = 1.7; 95% CI 1.1–2.7) and poor work effectiveness (OR = 2.6; 95% CI 1.6–4.2).

To test whether lowest quintile of SES or high combined JCQ/ERI might be effect modifiers for injury status, we retested the key multivariate models from Table 4 including the appropriate interaction terms. None yielded a P-value <0.40, consistent with a negligible interaction effect. As a sensitivity analysis for the three measures that did not include work status outcomes, we also re-estimated the multivariable models restricting the population to only those individuals employed at their original worksite (N = 556). Results were similar to that of the original models, with the exception of three covariates for the

Table 4. Adjusted injury case status and significant covariates for poor functioning at follow-up within the GROW Study Cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Lowest quartile SF-12 PCS (N = 664) OR (95% CI)</th>
<th>Lowest quartile work effectiveness (N = 621) OR (95% CI)</th>
<th>2+ lost workdays in past 4 weeks (N = 621) OR (95% CI)</th>
<th>Changed site or not working (N = 664) OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case (versus ref)</td>
<td>1.4 (0.9–2.2)</td>
<td>0.9 (0.6–1.4)</td>
<td>1.5 (0.9–2.4)</td>
<td></td>
</tr>
<tr>
<td>Age at follow-up (per 10 years)</td>
<td>1.3 (1.1–1.6)**</td>
<td>1.0 (0.8–1.2)</td>
<td>0.9 (0.8–1.1)</td>
<td>0.9 (0.7–1.2)</td>
</tr>
<tr>
<td>SES quintile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (scores 2–3)</td>
<td>2.0 (1.0–4.0)*</td>
<td>1.0 (0.5–2.2)</td>
<td>1.5 (0.7–2.9)</td>
<td>2.5 (1.1–5.8)*</td>
</tr>
<tr>
<td>2 (score 4)</td>
<td>1.2 (0.6–2.4)</td>
<td>1.7 (0.9–3.4)</td>
<td>1.4 (0.7–2.6)</td>
<td>1.6 (0.7–3.5)</td>
</tr>
<tr>
<td>3 (score 5)</td>
<td>1.5 (0.8–2.6)</td>
<td>1.4 (0.8–2.5)</td>
<td>1.5 (0.8–2.5)</td>
<td>1.5 (0.7–3.1)</td>
</tr>
<tr>
<td>4 (score 6)</td>
<td>1.8 (1.0–3.4)*</td>
<td>1.8 (1.0–3.2)</td>
<td>1.4 (0.8–2.4)</td>
<td>1.2 (0.6–2.6)</td>
</tr>
<tr>
<td>5 (scores 7–8) (ref)</td>
<td>1.0 (–)</td>
<td>1.0 (–)</td>
<td>1.0 (–)</td>
<td>1.0 (–)</td>
</tr>
<tr>
<td>Co-morbid hypertension</td>
<td>1.7 (1.1–2.6)*</td>
<td>1.5 (1.0–2.5)</td>
<td>1.2 (0.8–1.9)</td>
<td>0.7 (0.4–1.2)</td>
</tr>
<tr>
<td>Co-morbid diabetes</td>
<td>2.2 (1.0–4.7)*</td>
<td>0.6 (0.2–1.8)</td>
<td>0.8 (0.3–2.1)</td>
<td>1.9 (0.8–4.5)</td>
</tr>
<tr>
<td>Job strain/ERI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median: ERI ratio only</td>
<td>0.7 (0.4–1.4)</td>
<td>1.3 (0.7–2.3)</td>
<td>0.8 (0.4–1.5)</td>
<td>0.9 (0.4–1.9)</td>
</tr>
<tr>
<td>Above median: job strain only</td>
<td>1.5 (0.9–2.6)</td>
<td>1.3 (0.7–2.3)</td>
<td>1.1 (0.6–1.8)</td>
<td>1.4 (0.7–2.8)</td>
</tr>
<tr>
<td>Above median: both</td>
<td>1.7 (1.1–2.7)*</td>
<td>2.6 (1.6–4.2)**</td>
<td>1.0 (0.7–1.6)</td>
<td>1.2 (0.7–2.2)</td>
</tr>
<tr>
<td>Above median: neither (ref)</td>
<td>1.0 (–)</td>
<td>1.0 (–)</td>
<td>1.0 (–)</td>
<td>1.0 (–)</td>
</tr>
</tbody>
</table>

Predictors in multivariable models included age, gender and race/ethnicity, SES quintile, co-morbid hypertension and diabetes and combined effort-reward and job strain measure. Poor functioning defined as lowest functioning quartile over both assessments for the referents subsample for SF-12 PCS and work effectiveness. For lost workdays, poor functioning defined as >2 days lost in the past 4 weeks. N’s for outcomes applicable only to individuals working at follow-up are based on rounded number of individuals with an actual or imputed value of working at follow-up. N = 43 who were not working when cases and referents are pooled. SES = quintile SES, created by adding together one point each for higher levels of education and income shown in Supplementary Table 1 and grouping into quintiles, from lowest (1) to highest (5). Job strain/effort-reward ratio created by recording each variable into above-median (binary) equivalent and then testing each simultaneously to yield four categories. *P < 0.05; **P < 0.01; ***P < 0.001.
Discussion

Baseline deficits associated with incident WRMSDs were largely resolved by 2 years. Although we were able to show small, albeit statistically significant, injury-related differences at follow-up using continuous scales, clinically meaningful outcomes defined by poor status (as compared to better status) clearly demonstrated a marked attenuation over time of the baseline associations of WRMSD with PCS and body region-specific disability. Two risk factors were predictive of both poor health (PCS) and work status at follow-up: lowest quintile of SES and increased combined job strain/effort–reward imbalance at baseline.

These results suggest that the prognosis of occupational injury, of which WRMSDs are by far the most common, may involve subtle but clinically significant longer term deficits in health and work status. A growing body of literature exploring the relationship between exposure to adverse working conditions and health status underscores the need for further occupational research explicating their inter-relationship within the broader context of SES-associated health disparities [25]. This is especially relevant to recovery and work status in high-risk groups with the poorest working conditions [26–29].

When considered in light of the relatively sophisticated and well-utilized occupational health services at both study worksites, with all the care access advantages of these hospital settings, the results suggest that WRMSDs may carry a somewhat worse long-term prognosis—particularly for low-SES workers—than is generally acknowledged.

This study is limited insofar as the results may not be representative of workplaces in other industrial sectors, especially those with less access to occupational health care and rehabilitation and with potentially poorer employment conditions in terms of job strain/effort–reward imbalance. We also lacked data at multiple time points that might have allowed us to create a more detailed picture of the recovery trajectory. We recognize that social insurance schemas and related reporting biases, as they apply to work-related injuries, can vary internationally and intranationally. Nonetheless, this 2 year follow-up study should be relevant to the natural history of work injuries through whatever health system they may be tracked.

Further, this population had a fairly narrow SES range (all subjects were employed at study recruitment). Indeed, the ability to observe SES and work exposure effects in relation to health and work status under such limitations is all the more noteworthy. This invites serious consideration about the conceptualization and assessment of socio-economic status, as well as the embodiment of inequitable conditions that can potentially produce health inequities even within occupations [30]. Finally, a degree of selection bias could have been introduced in that study subjects who were reinterviewed had higher levels of education and differed by job group from those who were not. This theoretical limitation should have been largely offset by the relatively modest numbers lost to follow-up and by the use of multiple imputation techniques.

One implication of this analysis is that hospital workers with WRMSDs with relatively lower SES or higher job strain and ERI may be at risk of worse outcomes. This may justify targeted management and rehabilitation efforts for support service personnel and for health care aides who may be at ‘double jeopardy’ following incident WRMSDs.

Key points

- Baseline health deficits for most injured workers were largely resolved at follow-up.
- Socio-economic status and job strain and effort–reward imbalance effects on poor health and work status were evident at follow-up, even after accounting for injury status and other baseline factors.

Funding

U.S. National Institutes of Health (AR47798).

Acknowledgements

We are grateful to Laura Trupin, David Burian, Aurelio Font and Greg Ryan for their help with this paper as well as the labour unions and hospital management representatives for their assistance and cooperation.

Conflicts of interest

None declared.

References


