Expert Judgment study for Placement Ladder Bowtie

D. Kurowicka, R.Cooke, L. Goossens, B. Ale
Delft University of Technology , The Netherlands

ABSTRACT: Statistics of occupational accidents motivated a large project commissioned by The Ministry of Social Affairs and Employment in The Netherlands to reduce risks for workers. The model of an accident is represented by a so-called ”bowtie” diagram. To quantify bowties different sources of data were used. Quantification was mostly data driven, however as it is often the case the data are incomplete and then the expert judgment is employed. The first bowtie model has just been completed and this paper reports results of expert judgment study for this model.

1 INTRODUCTION

Despite efforts, according to the European Statistics on Accidents at Work (ESAW), every year in the 15 Member States of the EU about 5 million workers are victims of accidents at work leading to more than three days of absence from work; furthermore, about 5000 workers die. These statistics motivate a large project commissioned by The Ministry of Social Affairs and Employment in The Netherlands to reduce risks for workers. The model of an accident can be represented by a so-called ”bowtie” diagram. The term bowtie is used to refer to a structured model of the causes and effects of events. The bowtie model is found to be quite convenient in modelling the relevant accident scenarios. This project aims to establish bowties for (the most important) occupational accidents in the Netherlands and then combine them according to the exposure of the working population to specific hazards. 25 bowties has been identified. Bowties are selected according to importance and size of the problem. Falls are the leading cause of occupational fatality, as well as a major source of mortality (see Figure 1 Health and Safety Executive ”Statistics of Fatal Injuries 2003/04”]). Hence the first bowtie considered was the Placement Ladder Bowtie.

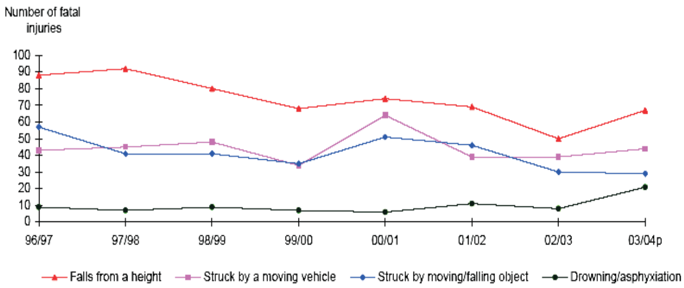

Figure 1: Number of fatal injuries to workers by kinds of accident 1996/97 to 2003/04 in the Great Britain.

2 PLACEMEN LADDER BOWTIE

Figure 2 shows the graphical representation of the Placement Ladder bowtie. The center event fall from placement ladder (F) is defined as a fall which results in death or serious physical and/or mental injury that has led to hospitalization or observation within 24 hours, as well as the suspicion of permanent physical or mental injury. Fall can be caused by failure of one of the primary safety barriers (PSBs): Ladder Strength (SR), Ladder Stability (SL) and User Stability (SU). Failure of one PSB is assumed to be sufficient for the fall. PSBs are influenced by support safety barriers (SSBs): Placement and Protection (PP), Right Ladder (RL) and Ability (AB). The SSBs in Figure 2 are influenced by Management but this problem will not be treated in this paper. On the right (consequence) side of the Placement Ladder bowtie we see four barriers that can influence Consequences (C) of the fall, that is, Ground (G), Height (H), Medical attention (M) and Age of the victim (A). In order to quantify this model with data, or if the data is not available by expert judgment, one must define precisely all barriers and events similarly to the definition of the fall. SR is defined as the provision of adequate strength to the ladder in order to carry the weight of the user and any additional loads; SL - provision of adequate stability to the ladder in order to provide a support to the user; SU - provision of adequate stability of the user on an otherwise strong and stable ladder. Each of these three PSBs can

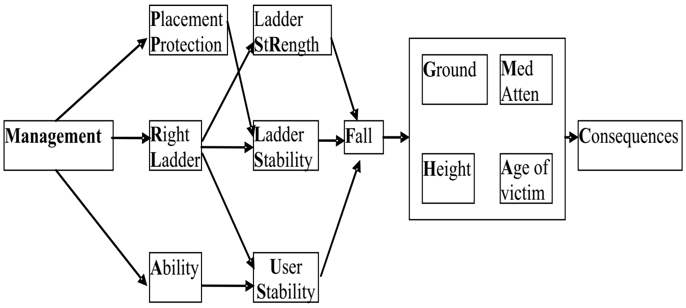

Figure 2: Placement Ladder bowtie model.

be in one of two possible states: success denoted by ”+” and Failure ”-”. The support safety barriers are all assumed to have two states ”+” and ”-”. Precise definitions follow:

1. Placement and Protection (PP) - a ladder is well Placed and Protected if all the following conditions are satisfied:
   - It is placed on the surface that is even and firm;
   - If placed at steep angle it is secured on the top;
   - If placed at too wide angle the measures are taken to prevent sliding;
   - If the ladder is long (7 or more meters) it is secured at the top and at the ground;
   - The measures are taken to prevent the ladder being hit by any object, an opening door and/or moving vehicles.
2. Right Ladder (RL) - a ladder is called the Right Ladder for the job if it satisfies all following conditions:
   - It is an industrial product for professional use;
   - It extends at least one meter above the standing step or the exit height;
   - It is tested and maintained;
   - It has the proper accessories for the use on slippery surface (as in the Arbobesluit) ;
   - Its steps and rungs are checked to be free from grease, oil, wet paint, mud, snow, ice, paper and other slippery materials.
3. Ability (AB) - includes all characteristics and conditions concerning the fitness and ability of the user on a strong and stable ladder. A person climbing a ladder is Able to do the job if he/she:
   - Climbs the ladder while facing it and keeping feet in the center of a ladder;
   - Does not climb it from the side or from above the top from another ladder;
   - Does not slide down the ladder;
   - Does not overreach or stand on the top rungs;
   - Does not move the ladder while standing on it;
   - Is well prepared to use the ladder that is has clean shoes from debris and slipper materials;
   - Is not under influence of alcohol, drugs, medication;
   - Does not feel sick, dizzy.

On the right hand side of this model we have Height of fall (H) with two possible outcomes:

$$H1 : h < 5m;\ H2 : h > 5m;$$

Type of Surface (G) with two possible states

$$G1 :\ Soft\ ground; G2 :\ Hard\ ground;$$

Medical Attention (M) with two states:

$$M1 :\ Prompt;\ M2 : Delayed;$$

Age of Fallen Person (A): with two states:

$$A1 : Age < 50\ years;\ A2 : Age > 50.$$

And finally three levels of consequences will be used: Consequences (C) with three states:

$$C3 : Death; C2 : Permanent\ Disability;$$

$$C1 : Recoverable\ Injury.$$

We want to calculate probability of fall given failure of one of SSBs that in principle can be influenced by management actions. This can lead to discovering ways to reduce risk and consequences of a fall.

Using elementary probability rules, one can express probability of fall given different outcomes of support safety barriers in terms of:

- probability of support safety barriers given one of the primary safety barrier lost and fall;
- probability of one of the primary safety barrier loss given fall;
- probability of fall;
- unconditional probability of different combinations of support safety barriers;

The first two, as well as the number of accidents necessary to find probability of fall can be obtained from data. One must however deal with missing data. This will be treated in Section 4. The exposure that is required to calculate the probability of fall and unconditional probability of support barrier's failure have to be gathered from e.g. expert judgment (see Section 3). The right hand side of the bowtie can be quantified from data and this way one could incorporate the consequence part of the model (Section 6). Sections 5 and 7 show results for the Placement Ladder Bowtie. Finally the last section contains some conclusions.

3 EXPERT JUDGMENT ANALYSIS FOR PLACEMENT LADDER BOWTIE

The performance-based structured expert judgment methodology has been applied in many risk and reliability studies (Cooke 1991). Experts are treated as another source of data. The model for combining expert judgments is called "classical" because it resembles in many ways classical statistical hypothesis testing. We first present briefly the classical model and then apply this technique to obtain variables of interest for Placement Ladder bowtie.

3.1 The classical performance based model for continuous variables

Experts state quantiles for their subjective distribution for each of the several uncertain quantities, say, 5%, 50%, 95%. The classical model constructs a weighted combination of expert probability assessments. These weights are based on two key performance measures, calibration and information, which are assessed on variables whose true values are known post hoc (though not known to the experts at the time of assessment). Calibration corresponds to statistical likelihood. In the language of statistics, this is the "p-value" at which we would reject the hypothesis that a given experts' probabilistic statements are true. Thus, low values for the calibration score (near zero) indicate low support for the hypothesis that the experts' probability statements are accurate; high values (near one) indicate high support for this hypothesis. Information or informativeness measures the degree to which the experts' distributions are concentrated. The weight obtained by expert is proportional to the product of his/her calibration and information scores, if the calibration scores exceed a "significance level" cutoff, which may be found by optimization.

The classical model computes "performance based" weighted combinations, but also uses the performance measures to assess the quality of other combinations. In particular, the performance of the equal weight combination is assessed. Generally, the combination exhibiting the best calibration and informativeness is recommended (Cooke 1991).

Id	Calibr.	Mean relat Total	Mean relat realizatii	Numb real	UnNormaliz weight	Normalized weight
1	0.001102	1.671	1.895	10	0.002089	0.4175
2	5.60E-05	1.888	1.794	10	0	0
3	9.87E-05	2.518	1.089	10	0	0
4	1.07E-06	2.017	1.978	10	0	0
5	0.000599	1.253	0.9288	10	0.000556	0.1111
6	0.00131	2.351	1.801	10	0.002359	0.4714
7	4.08E-06	1.577	1.479	10	0	0
Global	0.2441	0.8732	0.8507	10	0.2077	
Item	0.2441	1.208	0.9751	10	0.238	
Equal	0.3006	0.4729	0.4591	10	0.138	

Table 1: Calibration and information scores for experts.

3.2 Expert Judgment for Placement Ladder bowtie

First the identification and selection of experts was done in consultation with project oversight. Seven experts were selected; they are people with different professions ranging from manager of window cleaning company, through researcher in the institute for the construction company, to Labor Inspector. The elicitation questionnaire was designed to obtain the data required for the quantification of the Placement ladder. The elicitation format document contained basic definitions, 12 questions to assess the target variables and 10 questions to evaluate performance (obtained form different data banks containing information about fall accidents as well as published journal articles). One example of such a question is given below.

Question 1
In a hospital in a Western country about 150 patients were given emergency treatment in the 1990's after falling from a ladder. What percentage of the cases was the result of instability of the ladder (the ladder slid or tipped sideways)?

Experts first participated in a training session and then were elicited individually. A typical elicitation took 2 hours. Each expert gave 5%, 50% and 95% quantiles for 22 uncertain quantities. Table 2 shows the calibration and information scores for the seven experts in this study. All computations are determined using Excalibur. Tables and figures presented in this report are also the output generated by this software. The first column gives the expert number; the second column gives the calibration score.

The information scores for all variables and for the seed variables are shown in columns 3 and 4 respectively. The last column gives the "unnormalized weight"; this is the product of columns 2 and 4. If this column were normalized and used to form weighted combinations, experts 1 and 6 would be most influential. The numbers in Table 2 show that our experts are poorly calibrated. The worst calibrated expert is expert #7, the calibration scores for experts #1, #6 are the best. Experts can be combined in few different ways. The equal weight decision maker (DM) is obtained by assigning equal weight to expert's densities. Performance based decision makers are formed by weighted combinations of experts, where the weights are based on the experts' performance. Two performance based decision makers are supported in the software Excalibur.

- The ”Global weight” decision maker using average information over all calibration variables and one set of weights for all items.
- The ”item weight” decision maker constructs weights for each item separately, using the experts' information scores for the given item, rather than the average information score.

In this study the items weights perform better than global so we present calculation only for item weight decision maker. In this case, all experts with a calibration score less than the significance level found by the optimization procedure are unweighted as reflected by the zero's in column 7. We see that the item weight decision maker is slightly less well calibrated than the equal weight DM (calibration score equal to 0.3006), but this is more than offset by the gain in information (information for equal weight DM was equal to 0.4638). We see that item weight DM has lower information than each expert individually but the loss of information is much smaller than in the case of equal weights. Robustness analysis addresses the question, to what extent the results of the study would be affected by loss of a single expert or calibration variable. We compare the ”perturbed decision maker” to the original by computing the relative information of the perturbed to the original decision maker. It was discovered that the robustness of the item weight decision maker is quite satisfactory.

3.3 Elicited results

For Placement Ladder Bowtie there were 13 questions of interest. We present below 12 questions and the results calculated based on item weight decision maker by taking median value from its distribution. The question 13 was concerned the right hand side of the Placement ladder bowtie. It will be discussed in Section 6

1. Given 100 people chosen randomly from the Dutch working population, who use a placement ladder regularly solely as a means of transport for their work, how many ladder missions will they perform in a random week?
   Median: 990.4
2. Given a randomly chosen mission what is its duration, provided the ladder is used solely as a means of transport?
   Median: 21.41 in seconds.
3. Given 100 people chosen randomly from the Dutch working population, who use a placement ladder regularly as a work place too for their work, how many ladder missions will they perform in a random week?
   Median: 4739
4. Given a randomly chosen mission what is its duration provided the ladder is used also as a work place?
   Median: 467.7 in seconds.
5. What is the percentage of ladder missions in which the ladder used was not the Right Ladder?
   Median: 16.64
6. What is the percentage of ladder missions in which the ladder was not correctly Placed and Protected?
   Median: 16.71
7. What is the percentage of ladder missions in which the user was not Able to do the job?
   Median: 7.025
8. What is the percentage of ladder missions with not Right Ladder in which the ladder was not correctly Placed and Protected?
   Median: 10.05
9. What is the percentage of ladder missions in with wrong Placement and Protection in which the user was not Able to do the job?
   Median: 4.103
10. What is the percentage of ladder missions with not Right Ladder in which the user was not Able to do the job?
    Median: 6.926
11. What is the percentage of ladder missions that resulted in a Fall provided the (placement) ladder is used solely as a means of transport?
    Median: 0.0002071
12. What is the percentage of ladder missions that resulted in a Fall provided the (placement) ladder is used also as a work place?
    Median: 0.0001051

4 MISSING DATA

Developments of bowtie models as well as their quantification were to be data driven. However the data are often incomplete; necessary data fields are often simply not filled in. We first investigate the following problem and then generalize this to the multidimensional case: Consider two uncertain events A and B. The outcome of these events will be denoted as 1 and 2 . The outcome of these events may also be not recorded; we will treat these as unknowns. We would like to find a joint distribution over events A and B given the information in Table 2. To find this distribution we must ”re-distribute” the unknowns over the available cells.

$A \setminus B$	1	2	Unknown B
1	n_{11}	n_{12}	u_{1+}
2	n_{21}	n_{22}	u_{2+}
Unknown A	u_{+1}	u_{+2}	u_{++}

Table 2: Information about events A and B. n_{ij} - number of observations of $A = i$ and $B = j$, $i, j = 1, 2$. u_{i+} - number of observations in which $A = i$ and outcome of event B was not noted etc.

Different methods can be suggested as a solution for this problem. We present here only two that proved to be the best.

4.1 Maximum likelihood method

The standard statistical method that can be applied in this setting is the method of maximum likelihood (ML). The data in Table 2 can be seen as multinomial samples with general patterns of missing data. Denoting $\theta = (p_{11}, p_{12}, p_{21}, p_{22})$, corresponding to the probabilities for the four cells, one can write the likelihood function of θ given information in the Table 2.

$$\begin{aligned} L(\theta|X) &= P(X|\theta) = \\ &= \prod_{i,j=1}^{2} p_{ij}^{n_{ij}} \cdot \prod_{i=1}^{2} (p_{1+} + p_{2+})^{u_{i+}} \\ &\quad \cdot \prod_{j=1}^{2} (p_{+1} + p_{+2})^{u_{+j}} \end{aligned} \tag{1}$$

where $X = \{n_{11}, n_{12}, n_{21}, n_{22}, u_{1+}, u_{2+}, u_{+1}, u_{+2}, u_{++}\}$ and p_{i+} and p_{+j} denote marginal distributions of θ. Notice that the term $(p_{11} + p_{12} + p_{21} + p_{22})^{u_{++}}$ plays no role as evidently $p_{11} + p_{12} + p_{21} + p_{22} = 1$. Thus, observations for which neither the A nor the B value are known do not influence the likelihood. For special cases when there is no missing data or there is missing data only for one variable the maximum likelihood estimators are simple and analytic solutions are available (see e.g (Little and Rubin 1987)). For the general case an iterative procedure for ML estimation is required. The EM algorithm is used for this purpose. The E step of the EM algorithm is defined as follows: let $p_{ij}^{(t)}$ denotes current estimate of p_{ij} then:

$$n_{ij}^{(t)} = n_{ij} + u_{i+} \frac{p_{ij}^{(t)}}{p_{i1}^{(t)} + p_{i2}^{(t)}} + u_{+j} \frac{p_{ij}^{(t)}}{p_{1j}^{(t)} + p_{2j}^{(t)}}.$$

The M step calculates new parameter estimates as

$$p_{ij}^{(t+1)} = n_{ij}^{(t)} / \sum_{i,j=1}^{2} n_{ij}^{(t)}.$$

This method is very simple, easy to implement. The version of this algorithm is implemented in S-Plus.

4.2 Bayesian method

Bayes' Rule indicates how observations change beliefs. Suppose that we are interested in some parameter, say θ. First we associate a probability distribution with θ representing our prior beliefs (the prior probability). We perform observations and calculate the updated probability of θ using Bayes' rule (so-called posterior probability). This can be expressed as:

$$p(\theta|X) = \frac{p(X|\theta)p(\theta)}{p(X)} = \frac{p(X|\theta)p(\theta)}{\int p(X|\theta)p(\theta)d\theta} \tag{2}$$

where:

- θ is the parameter of interest;
- X is the observed data;
- $p(X|\theta)$ is the likelihood function of θ, ($L(\theta|X)$);
- $p(\theta)$ is the probability of the parameters value θ given only the prior beliefs. It is called the prior probability;
- $p(\theta|X)$ is the probability of the parameter value θ given the prior beliefs and the observations. It is called the posterior probability of the parameters;
- $p(X)$ is the unconditional probability of the data.

This rule says how the prior probability is replaced by the posterior after getting the observed data.

Suppose X has four possible outcomes, $\{(1, 1), (1, 2), (2, 1), (2, 2)\}$. In other words, the outcomes of X correspond to the cells in Table 2. We let $\theta = (p_{11}, p_{12}, p_{21}, p_{22})$, correspond to the probabilities for the four cells. We take a prior distribution over θ as the uninformative Dirichlet (1,1,1,1). It follows that the likelihood of the data represented in Table 2 is given by (1). The posterior distribution is proportional to the likelihood times the prior $p(\theta)$. If we now take the expectation of, say p_{11} with respect

(-,-)	(-,+)	(+,-)	(+,+)
0.6131099	0.386456	0.000374	0.00007

Table 3: Distribution $(PP, RL|SL-, F-)$.

(-,-)	(-,+)	(+,-)	(+,+)
0.317714	0.672807	0.004739	0.004739

Table 4: Distribution $(AB, RL|SU-, F-)$.

to this posterior, we obtain the updated probability for the cell $(1, 1)$.

$$\hat{p}_{11} = E(p_{11}) = \int p_{11} L(\theta|X) p(\theta) d\theta \qquad (3)$$

We note that the Bayes method with incomplete data becomes intractable if the distributions are large. The likelihood (1) has 8 factors. If there are n possible values of A, then there are $2^n - 1$ possible states of knowledge about A. If there are k variables, with variable j taking n_j possible values, then the number of factors in the corresponding likelihood (1) is equal to $\prod_{j=1}^{k}(2^{n_j} - 1) - 1$.

In our implementation we have sampled 50,000 samples from four independent gamma distributed random variables with parameters (1,1). Each sample was divided by its sum to obtain sample from Dirichlet distribution. Then the likelihood function was calculated and expected values corresponding to (3) were obtained from samples.

4.3 Redistribution for Placement Ladder Bowtie

PP\RL	+	-	Un RL
+	0	0	0
-	81	121	255
Un PP	2	21	2

AB\RL	+	-	Un RL
+	0	0	0
-	35	16	156
Un AB	0	0	0

Figure 3: GISAI data for LHS of Placement Ladder Bowtie.

Using GISAI data presented in Figure 3 and information that Ladder Strength was lost in 26 scenarios and in all of these cases Right ladder was not chosen we are able to find distributions of $(RL|SR-, F-)$, $(PP, RL|SL-, F-)$ and $(AB, RL|SU-, F-)$. We use Bayesian redistribution and obtain results shown in Tables 3, 4 and 5.

The distribution of $(AB, RL|SU-, F-)$ may be surprising at first as it means that the probability of ability lost and wrong ladder given user stability failure and fall is significantly smaller than when ability is bad but ladder is right. Even if this feels counterintuitive we must follow data analysis results that

(-)	(+)
0.9642857	0.0357142

Table 5: Distribution $(RL|SR-, F-)$.

found 35 accidents with $(AB-, RL+)$ and only 16 accidents with $(AB-, RL-)$.

5 RESULTS FOR LHS OF THE PLACEMENT LADDER BOWTIE

The expert judgment results are used to calculate probabilities on the left hand side of the Placement Ladder bowtie and for the center event. We show now how they were combined with information obtained from data to calculate the probability of fall given different outcomes of support safety barriers. We start with the probability of fall denoted as P(F-) . It can be estimated as the number of falls divided by the number of missions (where a ”mission” is defined as complete event on a ladder hence climbing on it, performing the required job and getting down on it). The number of falls is known form the data (GISAI data) and it is equal to 715.

The ladder might be used as means of transport (only ascent and descent) as well as work place (ascent, working while standing on the ladder, and descent). Let us introduce the following notations:

- MT - number of missions per person per week (ladder as transport),
- MWP - number of missions per person per week (ladder as work place),
- T - number of people using ladder as transport,
- WP - number of people using ladder as work place.

The number of missions (MT and MWP) was estimated via expert judgment (Question 1 and 3 in subsection 3.3). The number of people using ladders (T and WP) was estimated based on questionnaires sent to employs working with ladders [WoRM]. The number of missions over the period from which GISAI data was collected (6 years and three months) is:

#missions = { MT*T+MWP*WP}*#weeks*#years

$$= \{9.904 * 183751 + 47.39 * 77373\} * 42 * 6.25$$

$$= 1.44E + 9.$$

Hence the probability of fall was estimated as

$$P(F-) = 715/(1.44E + 9) = 4.96E - 7.$$

Now we calculate probability of primary safety barrier loss given fall denoted as $P(SL-|F-)$, $P(SR-|F-)$, $P(SU-|F-)$. They can be estimated from data as number of scenarios leading to fall due to PSB- divided by number of scenarios. Hence

$$P(SL-|F-) = \frac{482}{715};\ P(SR-|F-) = \frac{26}{715};$$

$$P(SU-|F-) = \frac{207}{715}.$$

To find distributions

$$P_1 = P(AB, PP, RL|SU-, F-),$$

$$P_2 = P(AB, PP, RL|SR-, F-),$$

$$P_3 = P(AB, PP, RL|SL-, F-)$$

we notice that in the bowtie model presented in Figure 2 we have

- Ladder Strength is only influenced by Right Ladder,
- Ladder Stability is influenced by Placement and Protection and Right Ladder,
- User Stability is influenced by Right Ladder and Ability.

From above information we obtain that AB and PP are independent of RL given $SR-$ and $F-$ which we denote as: $\{AB, PP\} \perp RL|\{SR-, F-\}$. We also have that $AB \perp \{SR-, F-\}$ and $PP \perp \{SR-, F-\}$. Thus, given that ladder strength has failed, the failure of user ability is no more likely than it is in the general population of ladder missions, and similarly for PP. From this information we have that:

$$P_1 = P(RL|SR-, F-)P(AB)P(PP).$$

Similarly one can obtain that:

$$P_2 = P(PP, RL|SL-, F-)P(AB),$$

$$P_3 = P(AB, RL|SU-, F-)P(PP).$$

The probabilities $P(AB)$ and $P(PP)$ were assessed by experts (Questions 6 and 7 in subsection 3.3). The distributions $P(RL|SR-, F-)$, $P(PP, RL|SL-, F-)$ and $P(AB, RL|SU-, F-)$ are given in Tables 3, 4 and 5.

The joint probability (AB, PP, RL) had to be estimated with expert judgment. To create this distribution in principle eight questions would have to be asked to experts. To reduce number of these questions it is reasonable to make an assumption that PP and RL are conditionally independent given AB.

(-,-,-)	(-,-,+)	(-,+,-)	(-,+,+)
0.04045	0.00058	0.02881	0.00041
(+,-,-)	(+,-,+)	(+,+,-)	(+,+,+)
0.01317	0.11290	0.08397	0.71970

Table 6: Distribution (AB, PP, RL).

Then the number of questions can be reduced to six. This assumption can be motivated as follows: the user chooses the ladder and the way that it is placed; hence knowing AB, RL and PP don't contain any additional information about each other. Any dependence between RL and PP is caused by the ability of the user. Using the above assumption the joint distribution of (AB, PP, RL) was constructed:

$$P(AB, PP, RL) = \frac{P(RL, AB)P(PP, AB)}{P(AB)}$$

where $P(RL), P(PP), P(RL, PP), P(PP, AB)$ and $P(RL, AB)$ are calculated from the information in Questions 10-15 in subsection 3.3.

The distributions presented above are sufficient to quantify the left hand side of the Placement ladder bowtie.

6 RIGHT HAND SIDE OF THE PLACEMENT LADDER BOWTIE

Figure 4 shows the collected data for the RHS of the Placement Ladder Bowtie. Since the quality of the data for $C3$ (consequence 'Death') was very poor we have combined $C2$ and $C3$ and denoted this as $C1'$. The required distribution is a joint distribution of (C, H, A, M) given Fall (denoted simply as (C, H, A, M)). The maximum likelihood method (EM algorithm) in S-Plus was used to obtain distribution of (C, H, A, M) (see column 4 in Table 7).

We have used the Bayes redistribution method and obtained the following one dimensional marginal distributions

$$C1 = 0.843,\quad H1 = 0.893,$$

$$A1 = 0.711,\quad M1 = 0.658.$$

The vine-copula method (see e.g. (Kurowicka and Cooke 2006)) requires information about one dimensional margins and additional information about the dependence structure in form of (conditional) rank correlations. The marginal distributions can be taken from data. Correlations can be calculated from experts' assessment or form data.

To construct the distribution (C, H, A, M) the vine in Figure 5 is used. Correlations and conditional correlations were assessed by experts, and from data with Bayes redistribution.

➤ Death

A1	M1	M2	UnM
H1	0	0	4
H2	0	1	1
UnH	0	1	0

A2	M1	M2	UnM
H1	0	0	5
H2	0	0	0
UnH	0	0	1

➤ Permanent Disability

A1	M1	M2	UnM
H1	4	2	44
H2	0	0	9
UnH	1	0	10

A2	M1	M2	UnM
H1	2	1	19
H2	0	0	4
UnH	0	1	2

➤ Recoverable injury

A1	M1	M2	UnM
H1	29	10	296
H2	2	2	37
UnH	5	1	68

A2	M1	M2	UnM
H1	14	7	100
H2	0	1	5
UnH	1	1	21

Un A	M1	M2	UnM
H1	0	0	3
H2	0	0	0
UnH	0	0	0

Figure 4: GISAI data for RHS of Placement Ladder Bowtie.

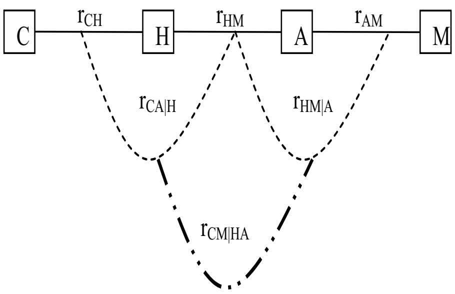

Figure 5: A D-vine on variables: consequence, height, age and medical attention.

Correlations from experts We incorporate the assumption that H, A, M are independent given fall, hence all correlations involving only these variables are zero (see Figure 6).

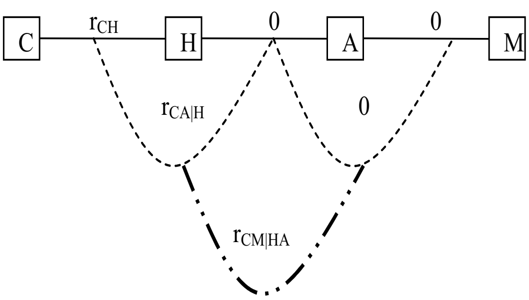

Figure 6: A D-vine on variables C,H,A,M where H,A,M are independent.

To construct the distribution (C, H, M, A) we must ask experts questions that allow us to infer values of the following rank correlations: r_{CH}, $r_{CA|H}$ and $r_{CM|HA}$. In the last question (question 13) the experts were asked to state the factor by which the probability of Death would change as the information $(H2)$, $(H2A2)$ and $(H2, A2, M2)$ became known. These factors could be combined with either equal or performance based weights (Cooke 1991), and multiplied by the marginal probability of Death to produce conditional probabilities. From these we could recover the rank correlations to be used when the joint occurrence of events is modelled with the diagonal band copula (details can be found in (Kurowicka et al. 2006)). Some of the experts gave factors so large, that the resulting conditional probabilities of death exceeded 1. These experts did not influence the performance based decision maker, but the occurrence of unrealistic assessments points to a weakness in the elicitation strategy. Another point is that the correlation method at present uses variables with two possible states. The variable Consequence takes three possible states. At the moment there is no corresponding theory that would give similar results for discrete variables with more then two states. Hence the queries to the experts were based on the binary variable C with states Death and Not Death and then use this correlation for the original distribution of consequences. This partition was chosen because Death seems ”more clear”. Recoverable injury, on the other hand, needs to be defined and assessing how specific factors may influence probability of this state could be more confusing. However, there are only 13 instances of Death in the GISAI data. For assessing correlations from data, we used the partition Recoverable C1 / Non-Recoverable injury (including Death) C2. The correlations obtained from the expert elicitation are given below.

$$r_{CH} = 0.048, \quad r_{CA|H} = 0.202, \quad r_{CM|HA} = 0.065.$$

Notice that in this case we have assumed that the conditional correlations are constant. We see that the experts do not think that the consequence Death is strongly influenced by Age, Height and Medical attention.

Correlations from data We have only thirteen cases for the consequence ”Death” and almost all of them with some unknowns. For consequence ”Recoverable injury”, on the other hand, the situation seems to be much better (see Figure 4). We have a large number of unknowns, but 603 scenarios gives a more representative sample. We can assume in the vine-copula that conditional correlations do not depend on the values of the conditioning variables (later on we will also use non constant rank correlations). We can therefore conditionalize on the variables with the highest counts in the data, and estimate the correlations to be used in the diagonal band for simulation with a vine from the data. However before we find correlation and conditional correlations the two dimensional and conditional two dimensional distributions must be obtained. We use Bayes redistribution method for this purpose. The results are shown below.

Distr	$(1,1)$	$(1',1)$	$(1,2)$	$(1',2)$
(C,H)	0.7615	0.1322	0.0805	0.0257

Distr	$(1,1)$	$(1',1)$	$(1,2)$	$(1',2)$
$(C,A\|H1)$	0.6686	0.0825	0.2439	0.0049
$(C,A\|H2)$	0.3905	0.5143	0.0571	0.0381

Distr	$(1,1)$	$(1',1)$	$(1,2)$	$(1',2)$
$(C,M\|H1A1)$	0.6305	0.0893	0.2266	0.0535
$(C,M\|H2A1)$	0.3744	0.0795	0.3901	0.1560
$(C,M\|H1A2)$	0.6443	0.0235	0.3149	0.0174
$(C,M\|H2A2)$	0.1886	0.2159	0.3787	0.2168

Conditionalizing on $H1$ and $A1$ and assuming that conditional correlations are constant we get

$$r_{CH} = 0.39, \; r_{CA|H} = -0.24, \; r_{CM|HA} = 0.30.$$

The negative correlation between consequence and age, given height, is counterintuitive at first sight. However, we may reflect that older workers are more experienced, perhaps less susceptible to pressure and less willing to take chances.

The assumption about constant conditional correlations can be weakened to allow changes of correlations depending on states of conditioning variables. In this case the following correlations were found.

$$\begin{aligned} r_{CH} &= 0.39, \\ r_{CA|H1} &= -0.24, \; r_{CA|H2} = -0.3, \\ r_{CM|H1A1} &= 0.30, \; r_{CM|H2A1} = 0.20, \\ r_{CM|H1A2} &= 0.21, \; r_{CM|H2A2} = -0.22. \end{aligned}$$

Notice that correlations didn't change dramatically. Only correlation $r_{CM|H2A2} = -0.22$ is surprising at first but examining counts of C and M given $H2A2$ (see (4)) one can see that this distribution was not well represented by the data.

$$\begin{array}{cccc} & C1 & C1' & UnC \\ M1 & 0 & 0 & 0 \\ M2 & 1 & 0 & 0 \\ UnM & 5 & 4 & \end{array} \quad (4)$$

The results obtained with vine method with constant and non constant rank correlations as well as redistribution obtained in S-Plus with the maximum likelihood method are shown in Table7.

(C,H,A,M)	v ncon	v con	S-Plus	Exp
p(1',1,1,1)	0.06288	0.07578	0.0584	0.0656
p(1',2,1,1)	0.01194	0.0139	0.0002	0.0079
p(1',1,2,1)	0.00994	0.0079	0.0225	0.0267
p(1',2,2,1)	0.00224	0.00166	0.0004	0.0032
p(1',1,1,2)	0.05236	0.0438	0.0310	0.0341
p(1',2,1,2)	0.00942	0.00832	0.0181	0.0041
p(1',1,2,2)	0.00726	0.00484	0.0200	0.0139
p(1',2,2,2)	0.00046	0.00086	0.0060	0.0017
p(1,1,1,1)	0.35446	0.40616	0.4260	0.3522
p(1,2,1,1)	0.03706	0.04206	0.0358	0.0422
p(1,1,2,1)	0.15814	0.15218	0.1317	0.1431
p(1,2,2,1)	0.01866	0.0176	0.0000	0.0171
p(1,1,1,2)	0.1666	0.146	0.1380	0.1830
p(1,2,1,2)	0.01724	0.0139	0.0327	0.0219
p(1,1,2,2)	0.08066	0.05834	0.0692	0.0744
p(1,2,2,2)	0.01068	0.0067	0.0099	0.0089

Table 7: Distribution of (C,H,A,M) obtained with vine method with constant rank correlations (col 3), vine method with non constant rank correlations (col 2) and with maximum likelihood method using EM algorithm in S-Plus (col 4) and vines with expert's correlations (col 5).

7 CALCULATIONS WITH PLACEMENT LADDER BOWTIE

We can calculate probability of fall give different combinations of support safety barriers.

One maybe interested to see how the probability of fall changes when the information about bad ability, wrong and not well placed and protected ladder becomes available. This corresponds to the following probability

$$P(F-|AB-,PP-,RL-) =$$

$$= \frac{P(AB-,PP-,RL-|F-) * P(F-)}{P(AB-,PP-,RL-)}$$

The probability of support safety barriers given fall can be calculated as

$$P(AB-,PP-,RL-|F-) =$$

$$= P(AB-,PP-,RL-|SU-,F-)P(SU-|F-)$$

$$+P(AB-,PP-,RL-|SL-,F-)P(SL-|F-)$$

$$+P(AB-,PP-,RL-|SR-,F-)P(SR-|F-)$$

Accommodating assumptions about distributions of support safety barriers given primary safety barrier lost and fall we get

$$P(AB-,PP-,RL-|F-) =$$

$$= P(AB-, RL - |SU-, F-)P(PP-)P(SU - |F-)$$

$$+P(PP-, RL - |SL-, F-)P(AB-)P(SL - |F-)$$

$$+P(RL - |SR-, F-)P(AB-)P(PP-)P(SR - |F-)$$

From the above we get

$$P(F - |AB-, PP-, RL-) = 5.50047E - 07$$

which is not very different from unconditional probability of fall. However if we calculate probability of fall given that all support safety barriers were in state "+", hence $P(F - |AB+, PP+, RL+)$ we get 1.51225E-09 which is two fold smaller than unconditional probability of fall.

To check which support safety barrier is the most influential given fall we can calculate

$$P(AB - |F-) = 0.3367,\ P(PP - |F-) = 0.7283$$

and

$$P(RL - |F-) = 0.5420.$$

The left hand side of the bowtie can be combined with the right hand side/consequence side to see how many deaths one can expect given different combinations of support safety barriers and fall.

8 CONCLUSIONS

Expert judgement methods are recognized as additional source of data in situations where other types of data are not available. This is in general not a "cheap" source of data. It requires the choice of good experts, representing diverse points of view, preparation of good seed questions as well as questions of interest. Moreover the experts usually participate in training session during which they learn about procedures used in an elicitation and about techniques to combine their opinions. Experts are interviewed separately by team of two people. Average interview time in elicitation for Placement ladder bowtie was 2 hours.

The Placement ladder bowtie was only one of the models concerned with falls. 25 categories of accidents were recognized in the WORM project. Quantification of so many models with expert judgment will be cumbersome. Of course accidents can be grouped such that the same experts can be used to quantify related models. However it was recognized that few panels of experts will be necessary.

As Placement Ladder bowtie was a pilot model to test techniques that will be use in other bowties later in the project, the question was posed whether it is possible to choose the best experts during training session and use their weighted combinations on questions of interest. This would significantly reduce time spent on the expert judgment exercise.

It turned out the the policy of choosing the best expert from the training and using only this expert for the elicitation would have resulted in unacceptably poor performance. Suffices to say that the best expert in the training followed a new heuristic in the elicitation which had been developed by him after the training. He found data from the UK for construction, and applied this to answer the seed questions, with very high confidence. The elicitor recognized the danger in this heuristic but was unable to dissuade this expert without violating the independence of the expert. This heuristic was not successful, and this expert received a low score in the elicitation. This underscores the reasons for using multiple experts.

Even if expert judgment methods are expensive they allow defendable quantification of models that otherwise would be have to be fed with numbers that seem reasonable.

REFERENCES

Cooke, R. (1991). *Experts in Uncertainty*. New York: Oxford University Press.

Kurowicka, D. and R. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley.

Kurowicka, D., R. Cooke, L. Goossens, and B. Ale (2006). Quantification with imperfect data from different sources. *Proc. of ESREL'2006, Safety and Reliability for Managing Risk 2*, 897–904.

Little, R. and D. Rubin (1987). *Statistical analysis with missing data*. Wiley.