Expectation of sickness absence duration: a review on statements and methods used in guidelines in Europe and North America

Wout E.L. de Boer¹, S. Mohsen Mousavi¹, George L. Delclos²,³, Fernando G. Benavides⁴, Mercedes Lorente⁴, Regina Kunz¹

¹ Swiss Academy of Insurance Medicine, University Hospital Basel, Basel, Switzerland
² The University of Texas School of Public Health, Houston, TX, USA
³ Centre for Research in Occupational Health, Pompeu Fabra University, Barcelona, Spain
⁴ Caisse Nationale de l’Assurance Maladie des Travailleurs Salariés (CNAMTS), Paris, France

Correspondence: Wout de Boer, Swiss Academy of Insurance Medicine, University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland, Tel: +41615565133, Fax: +41612655534, e-mail: Wout.deboer@usb.ch

Background: Certifying physicians play a key role in the management of sickness absence and are often provided with guidelines. Some of these guidelines contain statements on expected sickness absence duration, according to diagnosis. We were interested in exploring the evidence base of these statements. Methods: We identified guidelines through a survey of EUMASS members and a literature search of the Internet and PubMed. We extracted the statements and methods from the guidelines. We compared: diagnoses that were addressed, expected durations and development processes followed. Next, we presented our findings to the developers, to afford them an opportunity to comment and/or correct any misinterpretations. Results: We identified 4 guidelines from social insurance institutions (France, Serbia, Spain and Sweden) and 4 guidelines from private organisations (1 Netherlands, 3 US). Guidelines addressed between 63 and some 63000 health conditions (ICD 10 codes). Health conditions overlapped among guidelines. Direct comparison is hampered by differences in coding (ICD 9 or 10) and level of aggregation (three or four digit, clustering of diseases and treatment situations). Expectations about duration are defined as minimum, maximum, and optimum or mean or median and percentile distribution, stratified to age and work requirements. In a sample of 5 diagnoses we found overlap in expected duration but also differences. Guidelines are developed differently, pragmatic expert consensus being used most, supplemented with data on sickness absence from different registers, other guidelines and non-systematic literature reviews. The effectiveness of these guidelines has not yet been formally evaluated. Conclusions: Expectations about duration of sickness absence by diagnosis are expressed in several guidelines. The expectations are difficult to compare, their evidence base is unclear and their effectiveness needs to be established.

Introduction

Sickness absence, the behaviour of a worker who, for physical, psychological or social reasons stays away from work, has a major impact on workers, families, health care systems and economies. Insurance companies are particularly affected by the duration of sickness absence. Certifying professionals, frequently general practitioners, tend to dislike discussions with patients about expected duration of sickness absence as they feel uncertain about it.²–⁷ Both insurers and certifying professionals have expressed the need for evidence-based guidelines on duration of sickness absence.⁸

Several such guidelines exist and some address, for many health conditions, consequences for work, treatment, rehabilitation and expected duration of sickness absence.⁸,⁹ Some guidelines focus on work-related injuries (e.g. workers’ compensation systems), others address a single disease¹⁰ and still others provide guidance on several diseases. Application of these guidelines can impact certification and management of sick leave in the individual case, especially if they claim to be state-of-the-art and/or evidence-based.¹¹,¹² By serving as a ‘neutral’ authority, guidelines may narrow discussions between provider and patient, in particular regarding expected duration of the sickness absence.

Return to work depends on disease but also on the nature and circumstances of the work, personal (age, education, sex), cultural and organisational factors, including accessibility and quality of health care delivery.¹³ Given all these determinants, it seems challenging to develop guidelines that include statements about expected duration of sickness absence for multiple diseases.

To add to the body of knowledge, we compiled and examined guidelines on sickness absence duration that were in use in different countries in Europe and North America. Our research questions were:

- Which guidelines addressing sickness absence duration are in use in Europe and North America?
- How do these guidelines compare with respect to:
  - Diagnoses treated
  - Expected duration of sickness absence
  - Methods of development
  - Have these guidelines been evaluated or validated?

Methods

We included guidelines in use, in any language, that contain statements about expected duration of sickness absence from work for specific diagnoses; that were developed for use by sick leave certifying physicians; and that cover multiple diseases.

We examined those in Dutch, English, French and Spanish directly; those in other languages only if supported by experts of
the original research studies and reviews on observed duration of sickness absence; (ii) limited to sickness absence other than in work (e.g. school absences); (iii) only addressing work-related injuries and health conditions for use in workers’ compensation contexts; (iv) clinical guidelines on disease management and (v) guidelines that had been, but were no longer, in use.

Question 1: The guidelines we were looking for tend to be published nationally, not through scientific literature databases such as PubMed. Therefore we searched in three ways: Medline, experts and the internet. In 2013, we searched MEDLINE using the search term sick* AND certify* AND guid*. We also searched with MeSH headings identified from two relevant papers. In 2014, we approached experts from Europe, Canada and the US. We approached representatives of the European Union of Medicine in Assurance and Social Security (EUMASS, www.eumass.com). EUMASS unites insurance medicine experts from 20 European countries who work in organisations of health, life or social insurance (public or private). The national representatives are chief medical advisers of social insurance and comparable, all at leading positions and with a long experience in this field. We also contacted researchers/experts in Canada, the US and Spain. Via email, we asked these experts about guidelines that met our criteria. In both 2014 and 2015, we conducted Internet searches through two general search engines (Google and Duckduckgo) and one search engine on medical guidelines (tripdatabase.com) with the key words ‘disability guidelines,’ ‘disability duration’ ‘sickness absence duration’, ‘sick leave duration’ and ‘return to work.’ We examined each guideline that we found through any of these canals on it’s fitting to our inclusion criteria.

Question 2: In order to compare the included guidelines we used the published guidelines as sources and for further questions the developers of the guidelines. We extracted for each guideline the title, country, publishing organisation, publication date, target population and sample population (all workers or specific groups), target audience (sick leave certifying physicians and/or claims assessors) and the number of ICD Codes covered. For each guideline we extracted the basis on which diseases were selected for inclusion.

Across all guidelines, we compared expected duration as described in the guidelines. We selected five diagnoses that we considered representative for the spectrum of diseases in sickness absence of medium duration, converted to ICD-10 codes (adjustment disorder F43, depression F32-33, acute myocardial infarction I21, low back pain M54 and breast cancer C50). Regarding duration of sickness absence, we extracted: (i) reporting as statistical measures (percentiles, mean and median), as estimations (such as optimum) or otherwise; and (ii) whether subgroups were specified (e.g. age, sex, disease severity, complications or comorbidities, work requirements).

We examined all guidelines about information on their development. If expert consensus was used, we summarized expert credentials, and the method used to reach agreement. If data registries were used, we assessed the quality of summarized expert credentials, and the method used to reach consensus. If expert consensus was used to phrase expectations, we requested their review for accuracy and clarification.

Results

Guideline identification

We received 20 replies from the 23 countries we contacted. Four countries (France, Netherlands, Spain and Sweden) reported guidelines meeting our criteria.

Our internet searches identified three more guidelines: the American Medical Association (AMA) Guides to Evaluation of Work Ability and Return to Work, a web-based return-to-work toolkit (MD Guides or MDG, previously known as MD Advisor, henceforth US 1) and the Official Disability Guidelines (ODG, henceforth US 2). We excluded the AMA Guides as these are entirely based on the MDG.

Our MEDLINE search did not identify any additional guidelines.

Diagnoses treated

Guidelines addressed from 63 (French) up to 65000 (ODG) diagnoses, in ICD 9 (Spanish, US 1) or ICD 10 (Dutch, French, Swedish, US 2) codes. Diagnoses were sometimes split up according to treatment situation (e.g. after surgery or when using drugs).

Table 1 summarizes the selected guidelines and their characteristics.

 Statements on expected sickness absence

The five diseases we selected appeared with 59 ICD-10 codes in the guidelines (low back pain, 12 codes; adjustment disorder, 11; breast cancer, 9; depression, 14; acute coronary infarction, 13). For an overview of these codes in conditions of the low back, see table 2. The other health conditions are shown and compared in Supplementary Appendices 1–4. Of 12 codes of low back pain, only code M 54.5 (low back pain proper) was addressed in all guidelines, see table 2.

Six guidelines present estimations about expected duration of sickness absence, in terms of minimum, optimum or maximum; only Dutch does not. US 1, US 2 and Dutch present observations of duration. Dutch only includes cases that have lasted at least 42 days.

The US 1 and US 2 and French guidelines specify expected duration in relation to work demands (e.g. clerical/modified; manual; heavy manual); only US 2 defines these demands (Clerical/modified work: Lifting with knees (with a straight back, no stooping) not more than 5 lbs up to 3 times/h; squatting up to 4 times/h; standing or walking with a 5-min break at least every 20 minutes; sitting with a 5-min break every 30 min; no extremes of extension or flexion; no extremes of twisting; no climbing ladders; driving car only up to 2 h/day). In the Spanish guidelines the expected duration, a standard number of days, is stratified to age (<36, 36–55 and >55) with an adjustment factor, specific for each ICD 9 chapter (ranging from 0.70 to 1.30). The expected duration is further stratified by occupation, divided into 17 job groups such as clerks and construction workers (for all groups see Supplementary Appendix 1). The job groups were derived from the 2011 Spanish National Classification of Occupations. Each group of jobs gives an adjustment factor which is specific for each ICD 9 chapter (ranging from 0.63 to 1.39).

US 1 presents low back pain with observed sickness absence duration (41 905 cases). It suggests an optimum duration in low back pain of 1 (sedentary job) to 42 days (very heavy job) and presents an observed mean duration of 34 days. US 2 suggests 0–49 days (median, 17; mean, 28.8), whereas French recommends 1–35, Spanish 9–19 days and Swedish suggests maxima of 7–14 days.

Development of statements on expected duration of sickness absence

We asked contact persons from all guidelines about the development of the guidelines because none of the original publication described that in detail; see table 3. Diseases were selected either on prevalence in sick leave certification practice or following the ICD in its entirety.
Work requirements were included in a pragmatic fashion; none of the respondents referred to explicit rules in relating them to duration of absence. All guidelines relied on expert consensus, except the Dutch one which provides observational data. Procedures varied according to pathology; none of the respondents referred to systematic rules for reaching consensus. Spanish and Dutch used observational data from one source, Swedish used no primary data and the others used data from different sources; the manner of merging of data was not disclosed to us. All guidelines, except Dutch and Spanish referred to scientific literature, but not from systematic reviews.

**Evaluation of the statements on expected duration of sickness absence**

The guidelines French, Spanish and Swedish are developed and in use by insurance companies; US 1 and US 2 are in use with insurance companies and industries. Health professional are also expected to use the guidelines, in the individual case as indicative and supportive for decision making. It is unclear how the guidelines are used in practice; they might be used as evidence in cases where GP and insurance have different opinions. We found the following in the literature about the guidelines we identified. Skaner reported that 72% of responding GPs in Sweden said they used the Swedish guideline (but not how); 47% of these found the statements of expected duration problematic to use. Delclos et al. discussed an earlier version of the Spanish manual, finding the use of the observed mean as basis for the optimum, problematic. Nuckols and Ju compared the development of US 2 to the AGREE Criteria (the international tool for the assessment of practice guidelines http://www.agreetrust.org/) and found just over 80% correspondence. Ju also reports that it was uncertain if US 2’s search for evidence was comprehensive and if bias in the selection of articles was avoided.

**Discussion**

In a systematic survey among 24 countries in Europe and North America, we identified six guidelines for sick leave certification which we compared. These guidelines addressed from 63 up to 65 000 diseases, coded in ICD 9 or 10 at the 3 or 4 digit level, and contained expectations about sickness absence duration per diagnosis. The guidelines reviewed express these expectations in different ways (identification of disease, expression of the duration, stratification into subgroups of work requirements) and are therefore difficult to compare. For low back pain (M54.5), we found observed median durations ranging from 17 (US 2)–34 (US 1) days. It seems unlikely that differences in labour marker and health care services would account for these differences. Differences in the source population, data collection and merging might play a role here. The statements must often stem from expert consensus procedures that are not clearly described and, as far as we could find, not performed in a formalised manner. The evidence stems from eclectic literature searches and either unclear or unsystematic real life data. Dutch reproduces observed data and Spanish adjusts the mean ‘optimal’ duration in observed data by age and occupational factors. For low back pain maxima vary between 14 (Swedish) and 49 days (Spanish) with 19 (Spanish) 35 (French) and 42 days (US) in between. Differences in selection and use of the literature and in preferences of experts might play a role here.

Our study has several limitations. Other guidelines might exist, notably in countries other than those that responded to our request. From five diseases, common in sickness absence of medium duration, we could compare only low back pain, in all guidelines. It is possible that our findings would have been different for other diseases. We did not study how the expected (or recommended) durations correlate with ‘real world’ practice, as this was beyond the scope of the study; however, this might shed a different light on their impact. The way data are analysed and merged might play a role here.

The guideline developers used experts, scientific literature and real life data in a ‘pragmatic manner’. The differences in results are not negligible and there is no saying which guideline is the most right. All in all the statements seem to represent weak recommendations, based on a variable quality of evidence.

We asked ourselves if a better approach would be possible. Using one grid for many different diseases with each their own expression...
Table 2  Statements on sickness absence duration: disorders of the low back

<table>
<thead>
<tr>
<th>Health-condition</th>
<th>Expected number of days on sickleave</th>
<th>Observed number of days on sickleave</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>French (1) Optimum (range)</td>
<td>Spanish (2) Optimum (range)</td>
</tr>
<tr>
<td>Low Back Pain</td>
<td>M54.5</td>
<td>From 1 (sedentary work) to 35 (heavy-physical work)</td>
</tr>
<tr>
<td></td>
<td>M54.3</td>
<td>From 2 (sedentary work) to 35 (heavy-physical work)</td>
</tr>
<tr>
<td>Sciatica</td>
<td>M54.15 M54.16 M54.17</td>
<td>--</td>
</tr>
<tr>
<td>Radiculo-pathy</td>
<td>M54.89</td>
<td>--</td>
</tr>
<tr>
<td>Other Dorsalgia</td>
<td>M54.9</td>
<td>--</td>
</tr>
<tr>
<td>Dorsalgia unspecified</td>
<td>M54.26</td>
<td>From 21 (sedentary work) to 84 (heavy-physical work)</td>
</tr>
<tr>
<td>Other specified</td>
<td>M54.1.36 M54.1.37</td>
<td>From 21 (sedentary work) to 84 (heavy-physical work)</td>
</tr>
<tr>
<td>inter-vertebral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>disc displacement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other specified</td>
<td>M54.1.36 M54.1.37</td>
<td>From 21 (sedentary work) to 84 (heavy-physical work)</td>
</tr>
<tr>
<td>inter-vertebral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>disc degeneration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spondylolytic</td>
<td>S33.5</td>
<td>--</td>
</tr>
</tbody>
</table>

Legend: ‘--’ no data; n.r.: not reported; pts.: patients.

(1) French Optimum: the number of days in which the majority of employees is able to resume working, (sedentary to physically heavy work); adjust for: age and physical condition of patient, psychological factors, opportunities for workplace adaptation, employment and (exceptionally in this fiche repère) socioeconomic context.

(2) Spanish Optimum: standard number of days, adjusted for occupational- and age-correction coefficients.

(3) Swedish Maximum: for physically easy jobs (little lifting, bending and twisting), and for physically demanding jobs (much lifting, bending and twisting).

(4) US 1 Optimum: the window in time during which most employees will return without risk to themselves and others.

(5) US 2 Optimum: expert opinions of optimal physiological healing times.

(6) Dutch only includes observations of cases of over 42 days of sickness absence, no expectations.
and situations after treatment, in connection to different kinds of work requirements seems challenging. Vonk Noordegraaf et al. developed expectations for work resumption for one health condition (status after hysterectomy). They recruited expert physicians (gynaecologists, GPs and OPs) through medical boards and based recommendations on a literature review and a modified Delphi procedure. The expert panel judged 38 different work activities relevant for convalescence recommendations, which led to a refined system of recommendations. This illustrates that developing guidelines in a systematic manner for many diseases and tuned to differences in work requirements, would require a huge effort.

Spanish uses observed sick leave in a systematic manner. Respondents from French and Swedish stated that diagnoses in their registers of sickness absence are partly unreliable. Moreover, registers are partly inaccessible because of privacy and legal considerations and partly incomplete when it comes to work requirements. Prospective studies on sickness absence according to health condition and work requirements (and ideally interventions to promote return to work) would be needed to fill the gap of data. A challenge would be to study these in an internationally comparable way.

The guidelines may have more impact than their evidence base allows, especially since many are developed by or for insurance companies. Evaluations of the practical impact, or formal validation of the expectations, have not yet been carried out, but would appear to be urgently needed.

Scientifically the guidelines on expected sickness absence duration go largely unnoticed and are not well integrated into practice models. Sickness absence and its certification and promotion of return to work have been studied extensively over the past decades and found to be of a complex and sometimes controversial nature. We need a better understanding of these processes in order to be able to define the possible contribution of guidelines in the certification of sickness absence.

Conclusions

In several countries, certifying physicians are provided with guidelines that contain statements about expected duration of sickness absence for different health conditions. These statements seem to have a limited base of evidence and an unknown impact.

Improvement is possible but faces large challenges in organisation and resources.

Ethical approval

Under Swiss law on health research this study is exempt from review by an ethics committee.

Supplementary data

Supplementary data are available at EURPUB online.

Acknowledgements

The authors thank all respondents for their additional information.

Funding

This study was mainly funded by the Swiss academy of insurance medicine (ASIM), University Hospital Basel. ASIM is funded by donations from public insurers and private insurance companies. The present study was initiated and carried out by ASIM. Swiss insurance association SVV funded the comparison of statements about duration of sick leave, table 3, but had no say in the process or results of this comparison.

Conflicts of interest: M.L. works at the institution that developed the Fiches Répères and was co-author of several of them but none of the ones we compared.

Key points

- Guidelines on expectation of duration of sickness absence are provided by different providers
- These guidelines give partly different recommendations and their evidence base is unclear
- So far, the effectiveness of these guidelines has not been evaluated
References


