Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Vignette-based triage decisions in musculoskeletal care among physiotherapists, physiotherapists licensed for direct access, and physicians in Germany: An online cross- sectional survey

Lea Overmann
Association for Fascia Research
Sina Sophie Hess
Private Clinic Grüner Kranz
Andreas Brandl
TUM
Robert Schleip
Robert.schleip@tum.de

TUM

Research Article

Keywords: Direct access, Musculoskeletal disorders, Clinical vignettes, Triage decision-making, Physiotherapy screening

Posted Date: May 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9124701/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Safe management of musculoskeletal complaints requires recognizing red flags and making appropriate triage decisions, which is especially important in settings that consider expanded first-contact roles for physiotherapists, while routine statutory direct access remains limited in Germany.

Methods

We conducted an online cross-sectional survey in Germany (CHERRIES) using a standardized 12-vignette questionnaire. Participants (including physiotherapists, Heilpraktiker-licensed physiotherapists, and physicians) selected one of three management options for each vignette. The primary outcome was overall triage accuracy (the proportion of correct classifications); secondary outcomes included vignette-level response patterns and attitudes toward direct access.

Results

Data from 110 respondents were analyzed (70 physiotherapists, 21 Heilpraktiker-licensed physiotherapists, 19 physicians). Overall triage accuracy was low and similar across groups (35.6%, 43.23%, and 36.9%, respectively), with no statistically significant difference (one-way ANOVA, $F(2,107) = 0.66$, $p = 0.519$). Accuracy varied substantially across vignettes, with notable clustering of very low correct-classification rates in specific cases (particularly vignettes 7 and 10) across all professions. Support for direct access varied by group, with physicians exhibiting greater opposition.

Conclusions

Vignette-based triage accuracy was consistent across professional groups but highly dependent on individual cases, underscoring the importance of context-specific vignettes and clear definitions of referral thresholds for discussions of safe first-contact musculoskeletal care.

Background

The ongoing demographic shift toward an aging global population is expected to significantly increase the prevalence of severe diseases worldwide [1]. Severe pathologies are generally defined as conditions that are life-threatening, cause serious functional impairment, or require urgent medical intervention within 48 hours or up to 30 days, depending on the condition [2].

For clinicians managing musculoskeletal (MSK) complaints, safe care depends on recognizing warning signs (“red flags”), assessing the risk of serious underlying disease, and choosing an appropriate course of action—whether to continue MSK care, initiate MSK care alongside medical evaluation, or refer for urgent assessment [3]. In physiotherapy, this process is often referred to as screening: a structured clinical reasoning method that combines history, physical examination, and contextual details to determine whether a patient is suitable for physiotherapy or requires medical referral [4].

These competencies are often discussed in relation to direct access, defined as patient-initiated access to physiotherapy without prior contact, diagnosis, or referral from a physician [5–7]. World Physiotherapy describes physiotherapists as potential first-contact practitioners, and international promotion of direct access aims to improve access and reduce delays [8]. International evidence also supports direct access in MSK care: a systematic review found better satisfaction and outcomes, fewer visits, less imaging and medication use, fewer additional appointments, and no evidence of harm compared with physician-referred episodes [9]. However, arrangements for direct access vary significantly among countries. According to World Physiotherapy’s 2020 data, direct access to physiotherapy varies substantially across Europe. Full direct access is limited to a minority of countries, whereas private-access or self-pay models are more common. For example, broader direct access exists in the United Kingdom and the Netherlands, while in Germany and Italy, self-referral is mainly restricted to private practice; by contrast, Belgium does not permit routine direct access [8]. This international variation highlights a recurring safety concern: when physiotherapists serve as first-contact providers, triage and referral decisions must be accurate enough to minimize the risk of missing serious pathology [10]. In Germany, direct access to physiotherapy has not yet become a standard part of the healthcare system. MSK care still relies heavily on physician referrals and the legal rules for prescribing and delegated care [7]. A significant legal milestone was the Federal Administrative Court’s decision allowing physiotherapists—under specific conditions—to treat privately insured and self-paying patients without a physician’s prescription. In practice, this requires obtaining a sector-specific Heilpraktiker licence limited to physiotherapy (HP-licensed physiotherapists), which typically involves additional training and assessment related to first-contact responsibilities, including screening for serious pathology, recognising red flags and contraindications, and making timely medical referrals. This pathway allows restricted first-contact physiotherapy care, primarily for self-paying and privately insured patients, and does not correspond to unrestricted statutory direct access [4, 11–15]. At the same time, professional and political discussions are

increasingly focused on expanding physiotherapists' independence. This shift is driven by demographic changes, multimorbidity, and inefficiencies in referral-based pathways, including reforms such as the Blankoverordnung ("blank prescription"), in which physiotherapy is prescribed by a physician, but the physiotherapist is granted greater autonomy within defined boundaries to conduct the initial physiotherapeutic assessment, determine treatment frequency and duration, and select and adapt the specific interventions; in statutory care, this model currently applies only to defined shoulder-related indications [16, 17]. Empirically evaluating screening and referral decisions in routine care is complex; therefore, many studies use standardized clinical vignettes to examine triage decisions in MSK care, often distinguishing MSK presentations from non-critical and critical medical cases [10]. The vignette instrument developed by Jette and colleagues has been widely cited and was designed to test whether physiotherapists would choose physiotherapy management or medical referral in a simulated direct-access setting [10]. In German-speaking regions, the approach was translated and used in educational and research settings, including the German version developed by Beyerlein (2010) based on Jette et al. (2006) [10, 18]. More recently, countries without direct access continue to employ vignette designs to assess keep–refer decision-making: in Austria, a national survey found that physiotherapists classified MSK and non-critical cases more accurately than critical medical cases and interpreted findings as indicating targeted training needs [19]. Complementing performance assessments, methodological research has increasingly focused on developing and validating vignette sets for non-direct-access contexts through multiple, interdisciplinary steps [20].

Despite this increasing body of research, evidence comparing vignette-based triage decisions across professional groups relevant to MSK care in Germany remains limited. This gap is significant because Germany already incorporates first-contact-like arrangements in the self-pay sector through the sectoral Heilpraktiker pathway, while also debating broader role expansion and educational reform [12, 13, 15, 16]. Therefore, this cross-sectional study used a vignette-based questionnaire adapted from Jette et al. (2006), which has been previously applied in German-language research, to compare the accuracy of vignette-based triage among physiotherapists, HP-licensed physiotherapists, and physicians in Germany. It also describes response patterns at the case level across groups [10, 18]. By offering a snapshot of current triage decisions in a setting where professional autonomy is actively discussed but educational preparation remains diverse, the study aims to inform ongoing debates on safe first-contact practices and the integration of screening competencies into curricula within the German context [4, 19, 21].

Methods

Study design

This was an online, cross-sectional survey reported in accordance with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [22]. The study evaluated clinical decision-making using standardized case vignettes to compare triage and management decisions among professional groups involved in musculoskeletal care in Germany.

Ethics, Informed Consent, and Data Protection

Prior to the start of the study, an institutional ethics pre-assessment was completed in accordance with Hochschule Fresenius's procedures. Based on this assessment, a formal submission to the Ethics Committee of Hochschule Fresenius was not required for this anonymous survey study. Therefore, no ethics approval reference number was assigned. The study was conducted in accordance with ethical principles of the Declaration of Helsinki. Participation was voluntary and could be stopped at any time without consequences. Before the survey began, participants received written information about the study and had to actively give informed consent to participate (electronic consent). Data were collected anonymously; no personally identifying information was recorded, and responses could not be linked back to individuals.

Participants and Eligibility Criteria

The target population comprised three professional groups involved in musculoskeletal patient care in Germany: physiotherapists, HP-licensed physiotherapists, and physicians. Eligibility was assessed using self-report screening items embedded in the questionnaire. Participants were eligible if they met the minimum age requirement for their professional group (at least 18 years for physiotherapists and physicians; at least 25 years for HP-licensed physiotherapists) and met the professional qualification criteria for their respective group. No minimum duration of professional experience was required. Respondents were excluded if they reported a prolonged interruption of professional practice (> 4 years or, among physiotherapists, failure to obtain the required continuing education credits), if they had not been professionally active for more than 3 months, or if their professional licence had been revoked.

Recruitment and Survey Administration

The survey was conducted using EvaSys SurveyGrid, a web-based platform for survey programming and management. The questionnaire was shared via a link and QR code. Recruitment was carried out through multiple channels, including social media (Instagram, Facebook, WhatsApp) and direct outreach to practices, hospitals, and clinics. Efforts were made to recruit through professional organizations: physiotherapy

associations were contacted but did not share the survey with their members; sectoral Heilpraktiker associations supported distribution via their websites and social media by forwarding the link to their members. For physicians, recruitment via the national medical association was not feasible; therefore, the survey link was distributed directly to medical practices, hospitals, and clinics. Additionally, a paper version was used to help increase the sample size. Recruitment took place from January 1 to March 31, 2021.

A study flyer featuring a QR code and profession-specific invitation text was used for recruitment. Allocation to professional group and subgroup categories (qualification strata) was included in the study procedures to facilitate comparisons across groups.

Survey Instrument and Vignette Scoring

The questionnaire used in this study was not developed specifically for this study. It was based on the standardized 12-vignette instrument originally published by Jette et al. (2006) and on the German-language version previously translated and applied by Beyerlein et al. (2010) [10, 18].

For each vignette, participants selected exactly one of three management options (single-choice response format): (i) physiotherapy treatment without medical referral, (ii) physiotherapy treatment with additional medical referral or clarification, or (iii) no physiotherapy treatment with immediate medical referral. For analysis, each vignette had a predefined "correct" response category determined by its type. Five cases were categorized as musculoskeletal; responses indicating physiotherapy treatment (with or without additional medical clarification) were scored as correct, while "immediate medical referral" was scored as incorrect. For non-critical medical cases, the correct response was "physiotherapy treatment plus additional medical clarification." For critical medical cases, the correct response was "no physiotherapy treatment, immediate medical referral."

Outcomes and variables

The primary outcome was participants' diagnostic and triage performance, measured as the proportion of correctly classified management decisions across the 12 case vignettes. Secondary outcomes included the proportions of correct responses categorized by vignette category (musculoskeletal, medical non-critical, and medical critical) and by professional group.

Potential explanatory variables were gathered at the beginning of the survey and included gender, age group, professional group (physiotherapist; HP-licensed physiotherapists; physician; other), educational background (e.g., state examination, bachelor's/master's degree in physiotherapy, sectoral Heilpraktiker qualification; medical specialty), and continuing education (manual therapy and orthopaedic manual therapy, either in training or completed with a certificate). Work-related characteristics included employment status (self-employed/employed/both), work setting (clinic/public institution, practice, both, or other), weekly hours spent on direct patient contact, and the percentage of orthopedic patients in routine practice (< 50% vs. ≥50%).

Attitudes toward direct access were measured using a single item ("I support direct access" vs. "I do not support direct access because …"), with an open-ended field for reasons provided by those who do not support direct access.

Sample size planning

An a priori sample size calculation was performed using G*Power (version 3.1.9.7) for a one-way analysis of variance across three groups. The calculation assumed a medium effect size ($f = 0.25$), a significance level of 5% ($\alpha = 0.05$), and a target power of 80%. The planned sample size was 129 participants (43 per group).

Statistical analysis

Survey data were exported from EvaSys SurveyGrid. Descriptive statistics, including frequencies, means, standard deviations, ranges, and medians, were calculated. Group differences in overall triage accuracy were analyzed using a one-way analysis of variance (ANOVA). The significance level was set at 5%. All statistical analyses were performed with SPSS Statistics Subscription (version 1.0.0.1508).

Results

Overall, 110 health professionals participated (70 physiotherapists, 21 HP-licensed physiotherapists, and 19 physicians). Across all professional groups, participants were predominantly of working age (25–63 years), and most reported working 31 or more hours per week. The physiotherapist and HP-licensed physiotherapists groups included a higher proportion of women, while physicians were more often men. Work settings varied by profession: physiotherapists and HP-licensed physiotherapists primarily worked in outpatient settings, whereas physicians more often worked in hospitals. Some respondents reported additional qualifications or specialties (e.g., manual therapy or osteopathy). Participant characteristics by professional group are summarized in Table 1.

Table 1
Participant Characteristics, Education, and Continuing Professional Development

Characteristic	Physiotherapists (N = 70)	HP-licensed PTs (N = 21)	Physicians (N = 19)
Sex (%) female / male	76 / 24	62 / 38	42 / 58
Age (%)	18–24: 17; 25–63: 81; >64: 2	25–63: 100	25–63: 90; >64: 10
Orthopaedic caseload (%) < 50 / >50)	30 / 70	29 / 71	74 / 36
Work setting (%) practice / clinic / other)	76 / 26 / 4	91 / 24 / 5	42 / 58 / 0
Weekly working hours (%) by category)	< 10: 11; 11–20: 9; 21–30: 20; 31–40: 46; >40: 14	< 10: 0; 11–20: 5; 21–30: 24; 31–40: 42; >40: 29	< 10: 0; 11–20: 10; 21–30: 11; 31–40: 47; >40: 32
Employment status	84.3% employed; 15.7% self-employed or mixed	57% self-employed; 14% mixed (self-/employed)	42% self-employed (practice); 58% employed (clinic)
Physiotherapy qualification	State examination: 51; Bachelor's degree: 17 not reported: 2	Not reported (all held sectoral Heilpraktiker licence restricted to physiotherapy)	Medical specialty distribution (see below)
Medical specialty (n)	—	—	General practice: 5; Anaesthesiology: 2; Cardiology: 1; Dentistry: 1; Internal medicine: 4; Not specified: 6
Manual Therapy (MT)	MT certificate: 20; MT in training: 8;	> 70% MT certificate; 10% MT in training (no further counts reported)	30% reported pursuing MT certificate
Additional qualifications (reported)	Manual lymph drainage (> 13 participants); plus PNF/Bobath; occasional taping, CMD, exercise therapy	Osteopathy; PNF/Bobath/neurophysiology; prior nursing; sports physiotherapy	Not Reported

Legend: Values are presented as a percentage of each professional group unless stated otherwise; counts (n) are shown where available. HP-licensed PTs = physiotherapists holding a Heilpraktiker licence restricted to physiotherapy; MT = manual therapy; OMT = orthopedic manual therapy; MLD = manual lymph drainage; PNF = proprioceptive neuromuscular facilitation; CMD = craniomandibular dysfunction. 89Percentages may not sum to 100% due to rounding or multiple-response categories.

Primary Outcome: Vignette Triage Accuracy

Overall triage accuracy was low and similar across the three professional groups. Physiotherapists correctly classified 35.6% of vignette decisions, HP-licensed physiotherapists 43.23%, and physicians 36.9% (Table 2). There was no statistically significant difference between groups (one-way ANOVA, $F(2,107) = 0.66$, $p = 0.519$).

Case-level Results

Correct-response rates varied significantly across vignettes and exhibited clustering of notably low performance in specific cases across all professional groups (Table 2). Vignette 7 showed very low correct classification across all groups (physicians: 0%; physiotherapists: 4.3%; HP-licensed physiotherapists: 4.8%). Vignette 10 also showed low accuracy across groups (physicians: 15.8%; physiotherapists: 8.6%; HP-licensed physiotherapists: 4.8%). In contrast, vignette 5 was correctly classified by all physicians (100%) and by approximately half of participants in the other two groups (physiotherapists, 54.3%; HP-licensed physiotherapists, 57.1%). Vignettes with the largest between-group differences included vignette 12 (physicians 73.3% vs physiotherapists 45.7%) and vignette 11 (physicians 10.5% vs HP-licensed physiotherapists 42.9%).

As a descriptive sensitivity analysis, vignettes 7 and 10 were excluded because both showed very low accuracy across all three professional groups. After excluding these two cases, the overall pattern of between-group differences remained small.

Table 2
Vignette-level triage accuracy

Vignettes	Physicians n/N (%), N = 19	Physiotherapists n/N (%), N = 70	HP-licensed PTs n/N (%), N = 21
1. Buttock claudication	4/19 (21.1)	35/70 (50.0)	9/21 (42.9)
2. Forefoot pain runner	7/19 (36.8)	22/70 (31.4)	8/21 (38.1)
3. Thoracic pain after slip	8/19 (42.1)	39/70 (55.7)	13/21 (61.9)
4. Chronic neck pain	8/19 (42.1)	41/70 (58.6)	14/21 (66.7)
5. Hot red knee	19/19 (100.0)	38/70 (54.3)	12/21 (57.1)
6. Bilateral knee pain	4/19 (21.1)	24/70 (34.3)	6/21 (28.6)
7. Wrist deformity after fall	0/19 (0.0)	3/70 (4.3)	1/21 (4.8)
8. Teen knee injury	6/19 (31.6)	37/70 (52.9)	14/21 (66.7)
9. Thoracolumbar back pain	7/19 (36.8)	29/70 (41.4)	8/21 (38.1)
10. Sternal chest pain	3/19 (15.8)	6/70 (8.6)	1/21 (4.8)
11. Right subcostal pain	2/19 (10.5)	23/70 (32.9)	9/21 (42.9)
12. Night thoracic pain	14/19 (73.3)	32/70 (45.7)	14/21 (66.7)

Legend: Values are number correct (n) out of group total (N) with percent in parentheses. HP-licensed PTs = physiotherapists holding a Heilpraktiker licence restricted to physiotherapy.

Attitudes Toward Direct Access

Support for direct access varied by professional groups. Among physicians, 53% opposed it; among HP-licensed physiotherapists, 91% supported it; and among physiotherapists, 79% supported it. Reported reasons for opposing direct access included concerns about insufficient differential diagnostic assessment, delayed recognition of conditions, and the belief that initial contact should remain under medical supervision.

Discussion

This cross-sectional survey is, to our knowledge, the first German study to compare vignette-based triage decisions among three professional groups—physiotherapists, HP-licensed physiotherapists, and physicians—using a standardized vignette tool. Overall performance was low and similar across groups (around 35–42% correct), with no statistically significant differences between professions (one-way ANOVA, $F(2,107) = 0.66$, $p = 0.519$). However, the distribution of responses across individual cases indicates that the overall score was mainly influenced by a small number of problematic vignettes—particularly vignettes 7 and 10—suggesting that the low overall accuracy may reflect a mismatch between the vignette instrument and the German clinical context rather than uniformly poor clinical reasoning among participants. Excluding the two outlier vignettes did not materially alter the descriptive pattern across professions, suggesting that these cases influenced the absolute scores more strongly than the relative group ranking.

Comparison with Previous Literature: Shared 12-Vignette Instrument

The 12-vignette questionnaire, based on Jette et al. (2006), is widely used internationally to assess how physiotherapists make management and triage decisions for both musculoskeletal and medical (non-critical and critical) cases [10, 19, 23]. Because the same vignette framework and response options are used, the results are comparable across countries, reducing concerns that differences are attributable to differing case sets.

In the Danish dataset, overall accuracy was relatively low across categories (musculoskeletal: 42%; non-critical medical: 37%; critical medical: 34%) [23]. At the vignette level, the “woman with pain around the sternum” (vignette 10) remained notable: 55% chose “no PT and GP referral,” 39% selected “PT but encourage GP contact,” and only 6% went with “PT only,” resulting in a correctness of 45% (Budtz et al., 2021). In the Austrian national survey, overall performance was significantly higher (average correctness: 70.5% musculoskeletal; 79.5% non-critical medical; 53.5% critical medical; overall mean 69.2%) [19]. However, vignette 10 again stood out as a clearly low-scoring case, with only 19.0% correct, despite most participants choosing referral (81.0% “refer”) [19].

Our results mirror this cross-country outlier pattern and show that vignette 10 is not only challenging internationally but also particularly difficult in the German sample: for case 10, the correct rate was 8.6% among physiotherapists, 4.8% among HP-licensed physiotherapists, and 15.8% among physicians (present study). Likewise, case 7 was extremely low across our groups (0% among physicians; 4.3% among physiotherapists; 4.8% among HP-licensed physiotherapists). Taken together, the repeated “vignette 10 problem” across Denmark, Austria, and our study strongly

suggests that this vignette is highly sensitive to ambiguity and referral thresholds embedded in the scoring key, rather than reflecting a Germany-specific deficit [19, 23]. This strengthens the need to critically evaluate context fit and reference-standard assumptions when interpreting low overall scores and when using the instrument to draw conclusions about readiness for expanded first-contact roles.

This interpretation is supported by evidence from German-speaking research using the same Jette-based questionnaire. Beyerlein et al. translated and used the instrument with the explicit aim of replicating the original method and reported substantially higher performance than we observed in our study, both overall and for several individual vignettes [18]. Their vignette-level results were exceptionally high for some cases (e.g., vignette 7), while other vignettes remained difficult (notably vignette 10) [18]. The significant gap between Beyerlein's results and ours—especially for vignette 7—raises an important measurement question: are we observing real differences in decision-making ability, or do the discrepancies primarily arise from differences in implementation, scoring assumptions, or context-specific decision thresholds?

The Physician Benchmark: How to Interpret 35–42%

While 35–42% may seem low, research using physician vignettes suggests these tools primarily measure standardized clinical decision-making within a specific scoring system rather than overall "general capability." In a validation study, Peabody et al. (2004) compared vignette scores with those of standardized patients (considered the gold standard) and found that vignettes can accurately estimate the quality of clinical practice (physicians scored 68% on vignettes when applying the same explicit criteria as standardized patients) [24]. Veloski et al. (2005) also argue that vignettes are efficient tools for isolating clinical decisions and controlling case-mix variation, providing a simple and cost-effective way to measure practice variation [25]. Taken together, this research supports two main points: (1) vignettes can be methodologically sound, and (2) very low percentages should prompt a review of the scoring assumptions and context fit (such as thresholds, role expectations, and pathway limitations), rather than immediate judgments of incompetence.

Vignette 10 as a Cross-Country Outlier

The most consistent and clinically important signal across studies is vignette 10: it appears to be systematically problematic across datasets, not just in our study. Even in Beyerlein's German implementation, vignette 10 yielded comparatively low correctness [18]. It was also observed that the Danish and Austrian datasets show this case as an especially pronounced low-scoring outlier, suggesting a recurring case-specific bias rather than a Germany-only effect. This cross-national convergence supports the interpretation that vignette 10 is highly sensitive to ambiguity and referral thresholds embedded in the reference standard. Clinically, chest pain scenarios—even when they contain musculoskeletal features—sit near high-stakes boundaries; risk-averse escalation can be defensible in many non-direct-access contexts. If the scoring key rewards confidence in a musculoskeletal pathway, cautious decisions may be penalized systematically.

Interpretation: Reference-Standard Mismatch and Context Effects

The significant difference between our results and Beyerlein's, despite using the same Jette-based instrument, indicates that vignette performance may be heavily influenced by implementation details and situational decision thresholds. Beyerlein also recognizes limitations that support this idea, such as the brevity of vignettes and the difficulties of cross-country comparisons due to differences in educational traditions and system pathways [18]. More generally, in settings without direct patient access, clinical reasoning and escalation decisions are often based on conservative risk management and trigger-based screening logic. When cases are presented as brief written vignettes without the ability to observe the patient or verify the history, ambiguity increases and may heighten threshold effects. In such cases, cautious rule-out reasoning can be considered "incorrect" if the reference standard assumes a higher level of certainty before referral.

Vignettes 7 and 10 demonstrate this point. Vignette 7 depicts a fall with signs consistent with a fracture, making referral for medical evaluation and imaging clinically justified. Vignette 10 shows acute chest pain with musculoskeletal features but possibly high perceived risk; in settings without direct access, conservative escalation might be preferred even when serious pathology is unlikely. If scoring rules are misaligned with local pathways and acceptable risk levels, these cases can disproportionately reduce overall performance and lead to misleading conclusions.

Strengths and Limitations

A strength of our work is the three-group comparison and vignette-level analysis, which revealed focused case-specific effects (notably vignettes 7 and 10). The main limitation is an imbalance in group sizes and limited power to detect between-group differences. As with all vignette designs, short written cases cannot replicate interactive examination, follow-up questions, and contextual cues that influence real-world referral decisions. Lastly, the discrepancy between our findings and Beyerlein's German implementation highlights a measurement validity issue that cannot be resolved from our dataset alone and requires direct investigation.

Clinical Implications

For primary care and interprofessional triage, the main implication is that unadapted vignette scores should not be used as standalone indicators of readiness for expanded first-contact roles. A more defensible approach is to adapt the vignette set and scoring key to the German context by explicitly defining what constitutes "appropriate management" and acceptable escalation thresholds under local responsibilities. However, the outliers in vignettes 7 and 10 should be regarded not only as a methodological weakness but also as clinically meaningful signals. These cases represent high-stakes presentations in which heightened sensitivity to possible fracture, serious thoracic pathology, and the need for timely

escalation is essential. Educationally, vignette-based training therefore remains valuable, but should place particular emphasis on such red-flag constellations and incorporate (i) context-relevant cases, (ii) explicit safety-netting language, and (iii) short rationale prompts (“Why did you choose this pathway?”) to distinguish knowledge gaps from risk-threshold decisions.

Future Research and Conclusion

Future studies should: (1) formally test the measurement equivalence of the Jette-based instrument in Germany; (2) establish an interdisciplinary German reference standard per vignette, especially for recurrent outliers like vignette 10; (3) evaluate whether including rationale prompts enhances interpretability; and (4) validate adapted vignettes against more comprehensive standards (e.g., simulated patients or structured clinical assessments). Overall, the consistent cross-country difficulty of vignette 10 and the significant discrepancy between our results and Beyerlein’s German replication underscore the need for context-specific adaptation and validation before drawing strong conclusions about professional competence from these scores.

Declarations

Author Contributions: writing – original draft: O.L., H.S.S. writing – review and editing: O.L., H.S.S., S. R., B.A. Conceptualization: O.L., H.S.S. Formal analysis: O.L. Investigation: O.L., H.S.S. Methodology: O.L., H.S.S. Project administration: O.L., H.S.S. Resources: O.L., H.S.S. Supervision: O.L., H.S.S. Visualization: O.L., H.S.S.

Competing interests: The authors declare no competing interests.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Consent to Publish: Not applicable.

Data availability: The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: The authors thank all participants for their time and contribution to the survey. The authors also thank Andreas Breitbach for his valuable preparation and support throughout the study.

References

1. Ferrari AJ, Santomauro DF, Aali A, Abate YH, Abbafati C, Abbastabar H, Abd ElHafeez S, Abdelmasseh M, Abd-Elsalam S, Abdollahi A, et al. Global burden of disease study 2021: systematic analysis for 371 diseases and injuries. Lancet. 2024;403:2133–61. 10.1016/S0140-6736(24)00757-8.
2. Peterson S, Heick J. Referral decision-making and care continuity in physical therapist practice. Phys Ther. 2023;103(5):pzad030. 10.1093/ptj/pzad030. PMID:37249536.
3. Storari L, Piai J, Zitti M, Raffaele G, Fiorentino F, Paciotti R, Garzonio F, Ganassin G, Dunning J, Rossettini G, et al. Standardized definition of red flags in musculoskeletal care: a comprehensive review of clinical practice guidelines. Med (Kaunas). 2025;61(6):1002. 10.3390/medicina61061002.
4. Lüdtke K, Grauel L, Laube D. Screening in der Physiotherapie: Das Flaggen-System. Warnsignale erkennen. 2nd ed. Stuttgart: Thieme; 2021.
5. Boissonnault WG, Ross MD. Physical therapists referring patients to physicians: a review of case reports and series. J Orthop Sports Phys Ther. 2012;42(5):446–54. 10.2519/jospt.2012.3890.
6. Piscitelli D, Furmanek MP, Meroni R, De Caro W, Pellicciari L. Direct access in physical therapy: a systematic review. Clin Ter. 2018;169(5):e249–60. 10.7417/CT.2018.2087.
7. Nast P, Wirz I, Scheermesser M, Schämann A. Das Modellvorhaben Physiotherapie gemäß § 63 Abs. 3b SGB V: Ein Forschungsprojekt der BIG direkt gesund und dem Bundesverband selbstständiger Physiotherapeuten IFK e.V. 2018 [cited 2021 Apr 5]. Available from: https://www.big-direkt.de/sites/default/files/2018-04/ZHAW%20Modellvorhaben%20Ergebnisbericht.pdf
8. World Physiotherapy. Direct access and self-referral / Country profile data (reference year 2020) [Internet]. 2021 [cited 2021 Apr 16]. Available from: https://world.physio/
9. Ojha HA, Snyder RS, Davenport TE. Direct access compared with referred physical therapy episodes of care: a systematic review. Phys Ther. 2014;94(1):14–30.
10. Jette DU, Ardleigh K, Chandler K, McShea L. Decision-making ability of physical therapists: physical therapy intervention or medical referral. Phys Ther. 2006;86(12):1619–29.
11. Bundesverband selbstständiger Physiotherapeuten. - IFK e.V. Berufsbild Physiotherapeut/in [Internet]. 2017 [cited 2021 Apr 16]. Available from: https://www.ifk.de/physiotherapie/berufsbild-physiotherapeut/

12. Bundesverwaltungsgericht (BVerwG). Urteil vom 26.08.2009–3 C 19.08 (Heilpraktikergesetz; sektorale Heilpraktikererlaubnis) [Internet]. 2009 [cited 2021 Apr 5]. Available from: https://www.bverwg.de/260809U3C19.08.0
13. VPT Magazin. Sektoraler Heilpraktiker Physiotherapie: Direktzugang, Abrechnung, Dokumentation und Haftung [Internet]. 2016 [cited 2021 Apr 5]. Available from: https://www.vpt.de/
14. Bundesverwaltungsgericht (BVerwG). Urteil vom 19.03.2013–3 C 16.12 (sektorale Heilpraktikererlaubnis/Erlaubnis beschränkt auf Physiotherapie) [Internet]. 2013 [cited 2021 Apr 5]. Available from: https://www.bverwg.de/190313U3C16.12.0
15. PHYSIO-DEUTSCHLAND. Studiengänge Physiotherapie in Deutschland (Stand. März 2021) / Beruf und Bildung [Internet]. 2021 [cited 2021 Apr 5]. Available from: https://www.physio-deutschland.de/fachkreise/beruf-und-bildung/ausbildung-und-studium.html
16. Konrad R. Der Direktzugang zur Physiotherapie in Deutschland: Chancen und Herausforderungen [master's thesis]. Osnabrück: Hochschule Osnabrück; 2017.
17. Deutscher Bundestag. Terminservice- und Versorgungsgesetz (TSVG) vom 6. Mai 2019 [Internet]. 2019 [cited 2021 Apr 5]. Available from: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl119s0646.pdf
18. Beyerlein C, Stieger A, von Wietersheim J. Erkennen deutsche Physiotherapeuten anhand von Fallbeispielen Risikofaktoren (Red Flags) für die Behandlung? (Do German physiotherapists recognise risk factors [red flags] by means of case studies?). Manuelle Ther. 2010;14:184–92. 10.1055/s-0029-1245907.
19. Janssen J, Lackenbauer W, Gasselich S, et al. The ability of Austrian registered physiotherapists to recognize serious pathology. BMC Prim Care. 2024;25:387. 10.1186/s12875-024-02634-8.
20. Lackenbauer W, Gasselich S, Lickel ME, Schabel L, Beikircher R, Keip C, Wieser M, Selfe J, Mazuquin B, Yeowell G, Janssen J. Development and validation of clinical vignettes to inform an educational intervention for physiotherapists to detect serious pathologies: a mixed-methods study. BMJ Open. 2025;15(8):e097107. 10.1136/bmjopen-2024-097107.
21. PHYSIO-DEUTSCHLAND. Professionalisierungsmatrix Physiotherapie (PromPT) / Akademisierung und Direktzugang [Internet]. 2017 [cited 2021 Apr 5]. Available from: https://www.physiodeutschland.de/fileadmin/data/bund/Dateien_oeffentlich/Beruf_und_Bildung/Ausbildung_und_Studium/PromPT_2017.pdf
22. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6(3):e34.
23. Budtz CR, Rønn-Smidt H, Thomsen JNL, Hansen RP, Christiansen DH. Primary care physiotherapists' ability to make correct management decisions - is there room for improvement? A mixed method study. BMC Fam Pract. 2021;22:196. 10.1186/s12875-021-01552-1.
24. Peabody JW, Luck J, Glassman P, Jain S, Hansen J, Spell M, Lee M. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. Ann Intern Med. 2004;141(10):771–80. 10.7326/0003-4819-141-10-200411160-00008.
25. Veloski J, Tai S, Evans AS, Nash DB. Clinical vignette-based surveys: a tool for assessing physician practice variation. Am J Med Qual. 2005;20(3):151–7. 10.1177/1062860605274520.